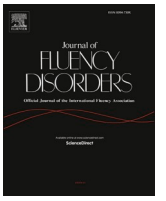
Journal of Fluency Disorders 87 (2026) 106194

Contents lists available at ScienceDirect

Journal of Fluency Disorders

journal homepage: www.elsevier.com/locate/jfludis



Speaking face-to-face with a virtual avatar to reduce anxiety in students who stutter: Tool development and pilot study results

Mathieu Delangle [a,b,1,*], Anne Moïse-Richard [a,c], Anne-Lise Leclercq [d], David Labbé [e], Stéphane Bouchard [f], Sheldon Andrews [e], Lucie Ménard [a,b]

[a] Marie Enfant Rehabilitation Centre, Sainte-Justine Hospital, Montreal, Canada
[b] Laboratory of Phonetics, University of Quebec in Montreal, Canada
[c] University of Montreal, Canada
[d] Research Unit for a Life-Course Perspective on Health and Education, Speech and Language Pathology Department, Université de Liège, Belgium
[e] École de Technologie Supérieure, Montréal, Canada
[f] Université du Québec en Outaouais, Canada

ARTICLE INFO

Keywords:
Virtual reality
Stuttering
Anxiety
School-age children
Exposure

ABSTRACT

Purpose: Speaking in class is challenging for students who stutter. Cognitive-behavioral therapy (CBT) with exposure in virtual reality (VR) emerges as a promising intervention for treating speaking anxiety in pediatric populations. This pilot study tested if real-time avatar-based VR can elicit anxiety responses while remaining acceptable to youth who stutter.
Method: Twelve students who stutter (aged 9–18) were randomly assigned to a single training session conducted either (1) in VR with a realistic avatar controlled live by their SLP, or (2) in role-play with their SLP, before facing a real actor. We assessed system acceptability, anxiety levels and perceived self-efficacy.
Results: The VR system was well accepted and elicited physiological arousal comparable to real-life interactions. Although participants reported experiencing less anxiety during VR, skin conductance level showed higher arousal; suggesting a divergence between the subjective report and physiological response. Finally, one training session (either in VR or with the SLP) did not produce gains in self-efficacy or decrease in anxiety related to the final real-actor conversation.
Conclusion: This study demonstrates evidence that the potential use of immersive VR could represent an acceptable and viable complementary strategy for SLP treatment, that could control exposure parameters while evoking physiological responses similar to real-life contexts. The differences between subjective and physiological measures suggest that VR is inducing anxiety responses differently than it was perceived. Further research could investigate the use of VR as anxiety interventions for students who stutter and should be explored across multi-session studies to understand their therapeutic effect.

* Correspondence to: Marie Enfant Rehabilitation Centre, Sainte-Justine Hospital, 5200 Rue Bélanger, Montréal, QC H1T 1C9, Canada.
E-mail address: mathieu.delangle@grenoble-inp.fr (M. Delangle).
[1] Present address: Grenoble INP, Grenoble University, 46 Av. Félix Viallet, 38000 Grenoble, France.

https://doi.org/10.1016/j.jfludis.2026.106194
Received 8 October 2025; Received in revised form 1 December 2025; Accepted 4 January 2026
Available online 13 January 2026

1. Introduction

1.1. Stuttering and anxiety in children and adolescents

Stuttering is characterized by disruptions in speech. However, it is the accompanying emotional distress and avoidance behaviors that often have the most pronounced effects (Gerlach-Houck et al., 2022). These effects may be further compounded by negative stigma (Johnson et al., 2024). School-aged children (Blood et al., 2011; Miluczenko et al., 2025) and adolescents who stutter (Erickson & Block, 2013) frequently face negative reactions from peers, which can lead to psychological problems, such as speech associated anxiety (Jones et al., 2021). A large body of research has concluded that children who stutter are more likely to report heightened anxiety versus fluent peers (Bernard et al., 2022). For example, Iverach et al. (2016) found that children who stutter were six times more likely to create social anxiety disorders compared to controls. In addition, social anxiety tends to increase over time, with peak during adolescence (Smith et al., 2014).

Stuttering can impact significantly academic progress and social participation (Johnson et al., 2023a) and school aged years can be a difficult time for children who stutter (Baxter et al., 2016). Cognitive behavioral therapy (CBT), particularly when watching graded exposure, tend to be effective at improving assertiveness and mental health with adults who stutter (Tse et al., 2023). CBT treatment helps adults to reduce their anxiety and their avoidance to the point they no longer meet diagnostic criteria of social phobia. However, there has been little empirical research on therapies designed to address the emotional impact of stuttering, including anxiety, for child and adolescent treatment (Johnson et al., 2023b; Kohmäscher et al., 2023).

1.1.1. Exposure exercises integrated into in cognitive-behavioral therapy

Exposure therapy, which involves repeatedly exposing oneself to stimuli that cause fear, is an integral part of cognitive behavioral therapy for anxiety disorders (Craske et al., 2014). Exposure exercises often involve real-life interactions, known as *in vivo* exposure. These exercises, such as interacting with a salesperson, closely mimic feared real-life situations and helps the person approach the moments they fear. However, logistical and temporal constraints, especially with public speaking tasks, can prevent repeated exposures from occurring (Bouchard et al., 2004, 2017). In addition, therapists have no control over *in vivo* treating context, which can elevate patient arousal and raise privacy issues (Bouchard et al., 2017; Stangier et al., 2003).

Alternative techniques, like imaginary scenarios and role-playing, address some limitations of *in vivo* exposure. For example, in treating stuttering anxiety, SLPs use role-playing to simulate stressful situations, with the therapist acting as another person (Murphy et al., 2007). This can help patients confront and desensitize to these moments, rather than avoid them (Menzies et al., 2009). It fosters the development of coping skills in controlled environments, with greater control over stimuli and variables. Role-play can also apply substantially more control over the stimuli and conditions in therapy. However, possibilities for generalization can be limited, as the therapist's office and the therapist act like the other partners are known and will make the scenarios much less like real-life *in vivo* challenges.

The effectiveness of exposure techniques is not limited to anxiety reduction but also involves changes in self-efficacy (Côté & Bouchard, 2009). Self-efficacy, defined as an individual's belief in their ability to effectively perform specific tasks (Bandura, 1986; Ornstein & Manning, 1985), is an important factor in skill acquisition and development. Previous studies have demonstrated that both adolescents (Manning, 1994) and adults (Ornstein & Manning, 1985) who stutter exhibit reduced levels of self-efficacy in communication relative to individuals who do not stutter. According to Bandura's social cognitive theory, self-efficacy is context dependent. Individuals can be low on self-efficacy for public speaking and high for skills that have no link, like performance on instruments and vice versa. Exposure to situations increase self-efficacy along with reduction of dysfunctional beliefs as most consistent predictors of positive outcomes (Côté & Bouchard, 2009). In this way, self-efficacy is an important goal together with anxiety reduction when practicing exposure techniques in individuals' who stutter.

1.1.2. Exposure exercises conducted in virtual reality

Virtual reality (VR) exposure in CBT holds potential as an alternative therapeutic technique for treating phobias and social anxiety in adults (Bouchard et al., 2017; Chesham et al., 2018; Horigome et al., 2020; Wechsler et al., 2019) and children (Beele et al., 2024; Man et al., 2024; Parrish et al., 2016; Wong Sarver et al., 2014). VR can generate physical, cognitive, and affective responses that may closely resemble real-world scenarios (Bell et al., 2020; Li et al., 2022; Parsons, 2015; Schöne et al., 2023). It allows exposure to anxiety-provoking situations in a controlled and safe simulated atmosphere.

Unlike role-play, VR based exposure allows patients to visualize stimuli or situations that they would have otherwise difficulty picturing. This allows the therapist total control of the images and situations evoked through the process and has the ability to duplicate many anxiety producing situations. In comparison to traditional *in vivo* exposure, VR helps the therapist provide contexts for repeated exposures in a secure and controlled atmosphere without having to set up real life experiences, as public speaking (Bell et al., 2020; Bouchard et al., 2007). Despite promising results for adults with anxiety disorders (Andersen et al., 2023; Anderson et al., 2013, 2016; Carl et al., 2019), research using VR for social anxiety in PWS, and particularly in young people who stutter, remains limited.

1.1.3. Virtual reality as a tool to elicit stress responses in children and adolescents who stutter

Limited research has explored the use of cognitive-behavioral therapy using exposure in virtual (CBT-EVR) for children and adolescents who stutter. The available studies in adults who stutter show promising results (Brundage et al., 2006, 2016; Brundage & Graap, 2004; Brundage & Hancock, 2015; Chard et al., 2023). For example, Brundage and Hancock (2015) indicated that a virtual audience elicited pre-performance anxiety similar to a live audience for adults who stutter. Similarly, Brundage et al. (2016) found

increased subjective anxiety levels when participants gave speeches to a virtual audience compared to an empty virtual room. These results are encouraging towards potentially utilizing virtual environments for therapy on anxiety for adults who stutter. Recent studies have also introduced VR scenarios for adolescents who stutter (Cecil et al., 2024; Tentu & Cecil, 2024), offering preliminary evidence regarding feasibility and acceptability. However, the development and evaluation of highly immersive, interactive VR environments specifically designed for school-aged children and adolescents who stutter remain at an early stage.

In Moïse-Richard et al. (2021), the authors addressed the possible ability of VR to elicit speaking anxiety in 10 children and adolescents who stutter (aged 9–17) through the use of speaking scenarios in front of a real audience, a neutral virtual classroom, a challenging virtual classroom, and a virtual room that was empty. The authors discovered that speaking in the virtual classroom elicited anxiety levels similar to scores on the real audience scenario. Similarly, self reported anxiety scores were more elevated within the context of the virtual classroom when compared to the empty room. These findings are promising when considering the possibility of using virtual environments to safely train young people who stutter to desensitize them to anxiety provoking situations.

On limitation Moïse-Richard et al. (2021) is that no physiological measures of arousal were included in the protocol. Participants only provided self-reported subjective measures to assess anxiety. The relationship between arousal and self-reported stress is not always straightforward (Weber et al., 2024). Self-reported measures may be influenced by individuals' self-awareness, social desirability or their ability to discern their own internal state (Chambers & Johnston, 2002; Dadds et al., 1998; Miers et al., 2011). In contrast, objective physiological measures provide a direct, continuous, and non-influenced response to anxiety stressors, readily able to measure autonomic nervous system activity that may be outside of awareness of objective report. Thus, pairing physiological and subjective measures can provide a more robust indication of anxiety responses to virtual environments in children and adolescents who stutter.

Moreover, the study by Moïse-Richard et al. (2021) focused on public speaking in front of a class. It does not fully capture the frequent, face-to-face interactions children typically encounter in everyday life. This limits generalization of results to other overarching social contexts. It might be useful to also examine whether other modes of conversational environments, such as a virtual environment simulating a one-on-one discussion with an avatar, can induce the anxious reactions generally observed in this type of situation.

Whereas previous studies targeted situations where participants made speeches to a group, such as oral classroom presentations (Brundage et al., 2016; Girondini et al., 2023, 2024; Glémarec et al., 2021; Moïse-Richard et al., 2021; Premkumar et al., 2021), most studies have not addressed the challenge of replicating everyday, naturalistic face-to-face interactions. In fact, these studies used pre-determined scenarios, which did not allow for the infinite number of possible responses generated in real time during social interactions. To advance the ecological validity of VR interventions for stuttering and social anxiety for training with potentially infinite social human reactions, we must create virtual environments that mimic the real-life conversational dynamics.

1.2. Simulating natural face-to-face interaction with a virtual avatar

The primary benefit of virtual reality is the potential to provide timely and personalized responses to participant behavior that emulate the active and reciprocal nature of social encounters in the real world (Son & Rubo, 2025). The immediate and personalized nature of the verbal and non-verbal feedback provided by an avatar that almost instantaneously mimics a real person's speech and movements, increases feelings of presence and ecological validity of the virtual experience. This aspect is crucial to obtain sincere emotional and behavioral responses. Naturalistic face-to-face interaction in virtual reality depends on two important aspects: dynamic interactions and graphical realism. These factors are related, but they support the realism and efficacy of the VR environments that simulate social interactions (Pan & Hamilton, 2018).

1.2.1. Graphical realism and nonverbal communication

Graphical qualities in the VR environment, including the appearance of the virtual avatar, contribute to the authenticity of events that simulate real-world public speaking interactions (Girondini et al., 2023). Avatars that can clearly express non-verbal cues, such as facial expressions, significantly improve the authenticity and interpretability of social interactions (Seymour et al., 2021). Furthermore, the clarity of these non-verbal signals is particularly important for participants, such as those with social anxiety, who are sensitive to social responses. Facial expressions play an essential role during social interactions, changing dynamically in response to situational demands and playing a crucial role in emotion recognition (Krumhuber et al., 2023). However, creating avatars capable of displaying in real-time a wide range of dynamic and realistic facial expressions is a significant challenge. Research has shown a tendency to use avatars with limited facial expressiveness, ranging from no expressions (Kampmann et al., 2016) to a predetermined set of static expressions (Glémarec et al., 2021; Moïse-Richard et al., 2021), such as just appearing happy or angry (Kiser et al., 2022), or only smiling, neutral or annoyed(Chittaro & Serafini, 2023). When these nonverbal expressions are blocked or overly static, emotional recognition performance can decrease significantly (Trautmann et al., 2009). Such limitations can affect the user's ability to recognize and interpret the avatar's emotions. However, emotional decoding of the person you're interacting with is a critical component of successfully managing the anxiety related to speaking in people who stutter. Indeed, this anxiety is close to social anxiety, characterized by the fear of negative evaluation by others, which is a central concept in contemporary cognitive models of social anxiety (Clark & Wells, 1995; Rapee & Heimberg, 1997) and is particularly relevant for children and adolescents who stutter (Iverach & Rapee, 2014).

1.2.2. Dynamic interactions and feedback

Natural interaction is dependent not only on verbal spoken, but mutual, dynamic responses. Dynamic verbal feedback, as well as

non-verbal (e.g., gestures, head movements, facial reactions) feedback are vital to participant engagement, and perceived credibility (Etienne et al., 2023). Both supportive or unsupportive feedback from audience members, has an impact on speaker anxiety (Kelly et al., 2007). Thus, it is crucial to include a variety of avatar behaviors to maintain the credibility of interaction (Glémarec, 2023). Nevertheless, real-time, fluid, interactive interactions in VR remains a challenge (El-Yamri et al., 2019). Some recent applications in speech-language pathology have begun to incorporate VR-based conversational partner (Leyns et al., 2025). However, current approaches often use scripted or timed reactions (Chard et al., 2023; Moïse-Richard et al., 2021). While such controlled behaviors are essential at this stage to ensure standardization and experimental reproducibility, these pre-determined responses can't replicate the spontaneous and naturalistic dynamics of human interaction.

1.2.3. The case for face-to-face interactions

The challenge is greater in face-to-face contexts, where participants are paying attention to only one interlocutor, in contrast to interactions in front of a group of avatars. Face-to-face interactions provide a richer nonverbal experience for participants, while also being the most engaging (and familiar) experience (Sprecher & Hampton, 2017). While technical solutions such as motion capture exist, they often require specialized equipment and remain difficult to implement in clinical settings.

To our knowledge, no previous study in stuttering has implemented a real-time VR environment to enable face-to-face interactions with photorealistic virtual avatar's verbal and non-verbal speech and behaviour. This research also uniquely integrates physiological and subjective measures of anxiety allowing direct comparisons between, *in virtuo* exposure (i.e., VR), with a SLP in role-play, and *in vivo* social situations.

1.3. Objectives

1.3.1. Objective 1

The primary objective of this study is to develop an affordable immersive VR system capable of facilitating dynamic, real-time interactions between a participant and a photorealistic virtual avatar, offering a level of flexibility and adaptability that goes beyond pre-defined scenarios and programmed responses, allowing for improved adaptation to individual behaviors and reactions. The first objective is also to evaluate user acceptance and sense of presence with this VR environment.

1.3.2. Objective 2

The secondary objective is to investigate whether real-time, dynamic interactions with a photorealistic virtual avatar, that displays both verbal and nonverbal responses, can elicit physiological and subjective anxiety responses in participants.

1.3.3. Objective 3

The third objective is to examine the potential added value of VR in terms of self-efficacy and anxiety reduction as compared to role-playing interactions with participants' own SLPs before facing *in vivo* speaking situations. To address these objectives, a randomized controlled pilot trial is conducted, in which young people who stutter have a speaking training session either (1) with a photorealistic avatar in virtual reality (*in virtuo*, experimental group) or (2) during role-playing with their own SLP (control group), before facing a teacher acted by an unknown actor in conditions close to real-life (*in vivo*, both groups).

2. Method

2.1. Participants

Thirteen children and adolescents who stutter (6 girls) were initially recruited to participate in the study, aged 9–18 years (M = 13.1 years, $SD = 3.0$). Participants were recruited from the Marie Enfant Rehabilitation Centre (CHU Sainte Justine) and the Raymond-Dewar Institute in Montreal, Quebec, Canada, and private SLP clinics. All participants had been professionally diagnosed as having stuttering by a certified SLP, and currently were in therapeutic treatment. To ensure a trusting relationship between the SLP and the patient, participants were required to have met their therapist for a minimum of three sessions. The study received approval from the Institutional Review Boards of CHU Sainte-Justine and the Raymond-Dewar Institute, and informed written consent was obtained from all parents and verbal consent for all participants. Each participant received a $50 gift card as compensation. One adolescent was excluded from the study because she had previously experienced a similar virtual environment during a television show recording to promote the present project. As a result, data from 12 participants were included in the final analysis.

Participants were randomly assigned to one of two groups: a control group, which completed the speaking training session with their SLP (in-person role-play), and an experimental group, which completed the session in the virtual reality environment. A clear randomized controlled pilot design was used to equally assign the participants to the two groups (n = 6 per group), using a random number generator. A Student's t-test confirmed age homogeneity between groups ($t(10) = -0.46$, $p = .653$). Additional details regarding the tasks and procedures for each group are provided in Section 2.3.

In order to characterize initial social anxiety and speaking confidence in amongst the participants, participants completed two assessments: the Personal Report of Confidence as a Speaker (PRCS) and the Liebowitz Social Anxiety Scale for Children and Adolescents (LSAS-CA). The French version of the short form of the Personal Report of Confidence as a Speaker Scale (PRCS) has 12 true/false statements about the participant's confidence in public speaking. Scores range from 0 to 12, with higher scores indicating lower confidence. The French short form of the PRCS has shown good internal reliability and structural validity, and convergent validity with

other social anxiety measures (Heeren et al., 2013). The French adaptation of the Liebowitz Social Anxiety Scale for Children and Adolescents (LSAS-CA, Masia Warner et al., 2003), developed by Schmits et al. (2014), consists of a total of 24 items, measuring fear and avoidance behaviors in social (12 items) and performance (12 alternates) situations. Participants rated their level of anxiety in each of the 24 scenarios of the LSAS call on a 4-point Likert scale, in addition to rating the level of avoidance in those situations on a 4-point scale. Research shown that the French version of the LSAS-CA has high internal consistency, and the structure validity is supported in a previous study by Schmits et al. (2014).

Group means and standard deviations for the PRCS and the LSAS-CA are listed in Table 1. Student's *t*-test (homogeneity of variances) indicates that there was no significant difference in the PRCS score between the control and experimental groups ($t(10) = 1.41$, $p = .188$). Mann–Whitney *U* test revealed no statistically significant difference in LSAS-CA scores between the experimental and control groups ($U = 14.00$, $p = .575$). Initial anxiety and avoidance behaviors in social and performance contexts were homogeneous between the two groups.

2.2. VR tool development

The developments aimed, on the one hand, to enable participants who stutter discuss in real time with a virtual teacher in a virtual classroom and, on the other hand, to enable speech therapists to interact in real time with their clients in the skin of a virtual teacher (Fig. 1). The developments were carried out using the Unreal Engine software (version 5.03) from Epic Games. The virtual classroom, which replicated the setup used in both the SLP role-play and *in vivo* conditions, was constructed using assets from the Quixel Bridge 3D asset library.

A photorealistic avatar of a virtual teacher was built using the MetaHuman platform. Both the avatar and environment were tailored for optimal rendering on the Meta Quest 2 Virtual Reality headset with high graphics quality and view independence. The headset was connected to a laptop running an Intel Core i7–12700H CPU at 2.30 GHz, 16 GB of RAM, and a GeForce RTX 3070i GPU.

The SLP controlled the avatar in real-time from a room adjacent to the experimental area. An iPhone 11 was placed in front of the therapist, utilizing the Live Link Face app, which tracked therapist's facial movements (lips, mouth, eyes, eyebrows, and head movements) in real-time (see Video S1). These data were streamed and mapped to the avatar's face, allowing the virtual teacher to produce a variety of facial expressions and nuanced nonverbal behaviors in real-time and dynamically. This enabled the avatar to express emotions (e.g., surprise, paying attention, or displeasure) using more naturalistic facial dynamics to increase the realism and responsiveness of the virtual interaction.

The therapist's voice is also altered using the Clownfish Voice Changer software. Since most Quebecois SLPs are female, we chose to simulate a male avatar to ensure that the avatar represented a person as different as possible from the SLPs of the children or adolescent who stutter. A male voice, approximately 40–50 years old, was used to match the appearance of the virtual teacher (see Fig. 2). The same standardized male voice effect was applied for all participants, ensuring a consistent auditory representation of the avatar throughout all sessions.

The SLP monitors in real-time, on a computer placed in front of her, what the participant sees in the VR headset and what he says to the teacher. This allows her to vary her spoken information based on her participant's actions and expressions of the participants, making for a responsive realistic exchange that involves verbal and nonverbal feedback. The avatar's body movements, including a slight rocking motion on the chair and arm movements, are animated and integrated using the Adobe Mixamo library (https://www.mixamo.com). These movements were designed to simulate natural, lifelike behavior and to avoid an unnatural static appearance. Importantly, these movements are neutral and do not represent specific positive or negative reactions.

While wearing the VR headset, the participant was completely immersed in the classroom environment, with a full 360-degree peripheral field of view (see Video S2). At the onset of the immersion, the participant will see a short black screen and hear children playing in the schoolyard next to the classroom. Then the immersive classroom will appear, with the teacher facing them and seated at their desk. The visual representation of the participant's hands and fingers are seamlessly mapped in real-time in the headset. In the experimentation room, the child is seated on a chair, in front of a desk that is an exact replica of their desk in the immersive classroom. The position and size of the virtual desk are the same as the desk in the experimentation room. The auditory (while sounds of the schoolyard), visual (view of the body of the participant in the headset) and tactile (the desk) feedback aim to increase the participant's sense of embodiment and presence in the virtual classroom simulation.

Table 1
Means and Standard Deviations of Personal Report of Confidence as a Speaker (PRCS), and of Liebowitz Social Anxiety Scale for Children and Adolescents (LSAS-CA) for control and experimental groups.

		PRCS	LSAS-CA Total	Fear	Avoidance	Social	Performance
Experimental group	M	6.17	57.00	28.50	28.50	30.50	26.50
	(SD)	(3.31)	(28.17)	(17.01)	(12.18)	(18.63)	(11.83)
Control group	M	3.83	68.67	34.83	33.83	36.67	32.00
	(SD)	(2.32)	(11.38)	(5.67)	(8.33)	(6.38)	(11.38)

Note. PRCS scores range from 0 to 12, with higher scores indicating less confidence in one's public speaking abilities. LSAS-CA scores range from 0 to 144 for the total score, and from 0 to 72 for each section (fear, avoidance, social, and performance), with higher scores indicating increased social anxiety.

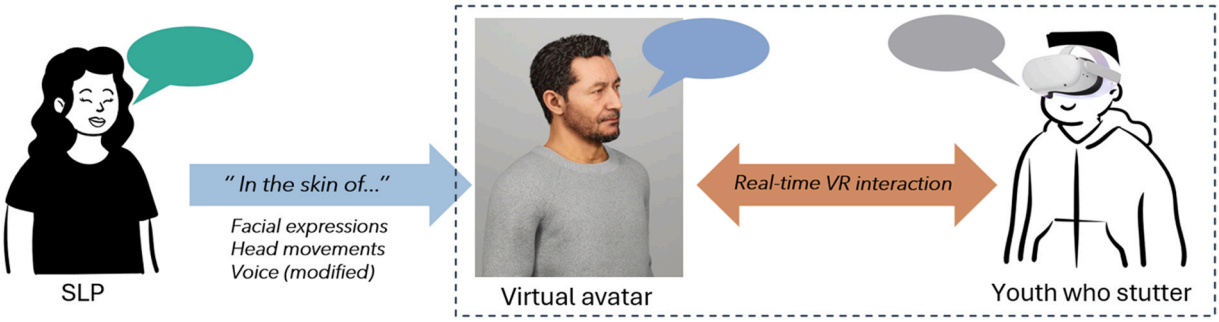


Fig. 1. Enabling real-time interaction between SLPs and clients through a virtual avatar framework.

Fig. 2. *In virtuo* experimental condition (experimental group): (Left) The virtual teacher; (Center) Participant seated at the desk, immersed using the VR headset and interacting in real-time with the virtual teacher; (Right) Virtual teacher in the virtual classroom seen by the participant through the VR headset.

2.3. Procedure

Participants completed two experimental sessions approximately two weeks apart (M = 12 days, SD = 11.3, range = 3–42 days). Most participants completed the two sessions within 7–12 days, with one shorter interval (3 days) and one longer interval (42 days) due to unforeseen scheduling constraints. The first session, lasting 45–60 min, involved either immersion in virtual reality (*in virtuo* condition, experimental group) or face-to-face interaction with their own SLP (control group). The second session lasted about 30 min and consisted of an *in vivo* condition, where participants from both groups interacted face-to-face with an unfamiliar actor-teacher. For the conversation tasks with the SLP and *in vivo*, recording started when the teacher entered the room and ended when the teacher closed the conversation. For the *in virtuo* conversation task, recording began when the simulation started and the teacher appeared in the VR headset and ended when the teacher closed the conversation. Each conversation lasted approximately five minutes (M = 4.7 min, SD = 1.6). Questionnaires and subjective measures were collected at specific time points, as illustrated Fig. 3.

Each session included three standardized phases, illustrated in Fig. 3: (1) Baseline, a neutral conversation designed to establish a physiological baseline; (2) Speech Preparation, a preparatory discussion addressing classroom accommodations; and (3) Conversation Task, an interaction with the teacher. Physiological anxiety measures were collected during each of these three phases:

1. **Baseline**: A three-minute conversation between the participant and an unfamiliar research team member (different from both the SLP and the actor-teacher), discussing neutral topics unrelated to stuttering or anxiety, was conducted at a desk identical to the experimental setup. This conversational baseline was selected to account for physiological activation naturally associated with speaking. Discussion topics may include hobbies, favorite music or books, and preferred travel destinations. The fact that our baseline is carried out with a person unknown to the participant, discussing topics that are positive for them, allows our design to evaluate the specific impact of the anxiety linked to having to explain and discuss one's stuttering in front of a new teacher. The SUDS score collected immediately after this phase served as the baseline reference in the presented analyses Section 4.
2. **Speech Preparation**: The participant had a three-minute preparatory discussion with the same research team member about classroom accommodations the participant might request regarding their stuttering. The research team member began by asking whether the participant already had ideas of accommodations they would like to discuss with the teacher. If the participant did not have suggestions, the research team member provided examples (e.g., not being called on to read aloud, giving oral presentations individually rather than in front of the class, or receiving support when being laughed at by peers) to help generate a short list of

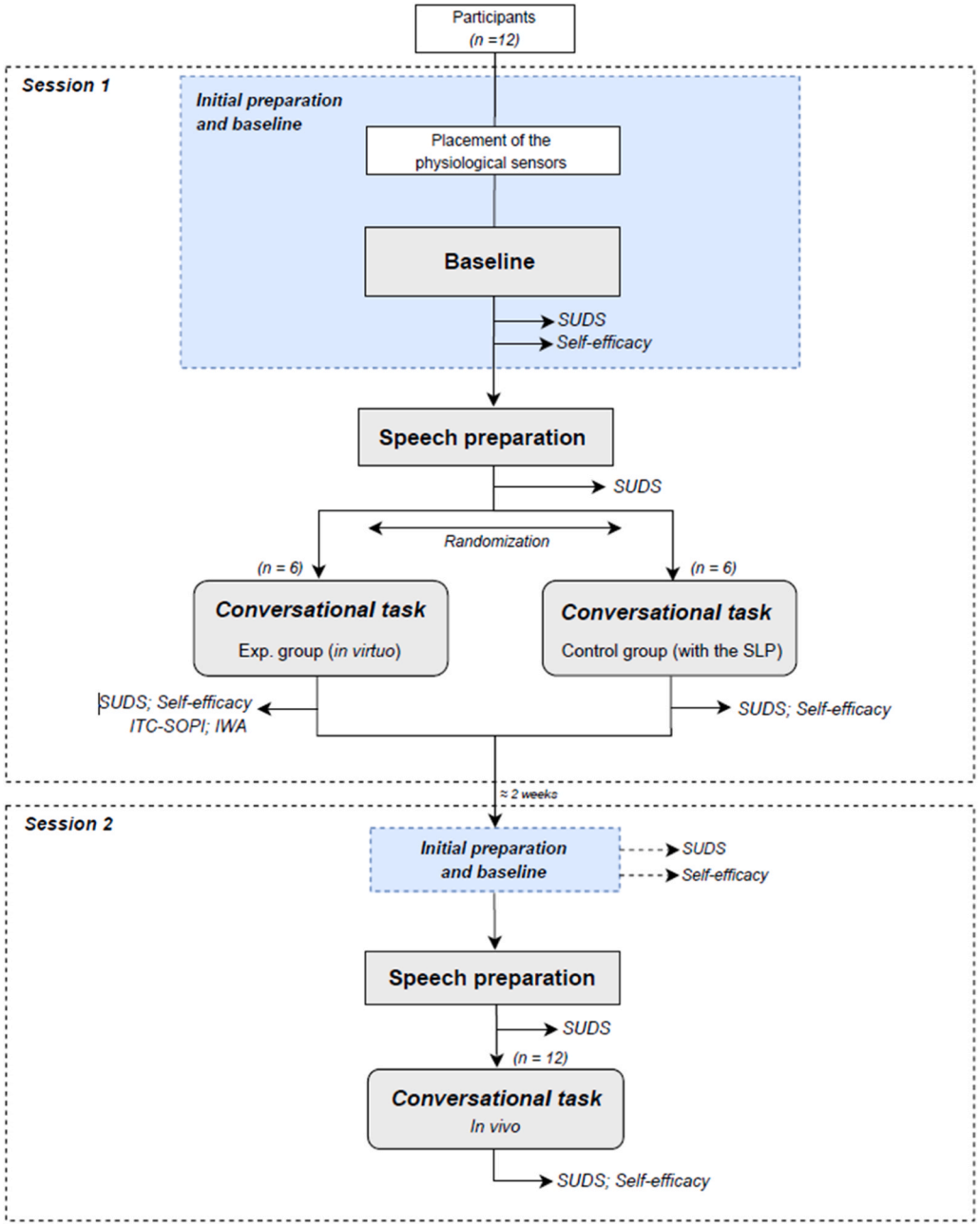


Fig. 3. Overview of the experiment timeline procedure and measures. Note. Sections shaded in gray (Baseline, Speech preparation, and Conversational task) indicate the phases during which physiological measurements were recorded. The "Initial preparation and baseline" section in Session 2 was identical to that in Session 1. SUDS: Subjective Units of Distress Scale; PRCS: Personal Report of Confidence as a Speaker; LSAS-CA: Liebowitz Social Anxiety Scale of Children and Adolescents; Self-efficacy (tailor-made questionnaire); ITC-SOPI: Sense of Presence Inventory; IWA: interaction with avatars (tailor-made questionnaire).

relevant accommodations. Although no standardized script was used, the prompts followed a consistent structure across participants to ensure procedural uniformity. The purpose of this phase is to prepare the participant for interacting with the teacher during the conversational task. Both physiological and subjective anxiety measures were collected during this phase, reflecting anxiety related to discussing accommodations and preparing conversational task with the team member.

3. **Conversation Task:** The participant had a five-minute ($M = 4.7$ min, $SD = 1.6$) interaction requiring to disclose their stuttering to the teacher and negotiate classroom accommodations. In all scenarios, the teacher discovered that the student stuttered. The teacher listened to the student's requests and engaged in dialogue, initially responding in a neutral and natural manner. As the interaction progressed, anxiety-provoking stimuli were introduced based on participant responses. These stimuli consisted of three clearly defined categories: frowning, yawning and uncomfortable verbal remarks such as "That's not fair to the other students" or "Why should I make special accommodations for you?". Each stimulus was introduced at least once during the interaction.

2.3.1. Conversation task during session 1

The first session took place in a dedicated room designed to accommodate both the SLP office setup and the VR simulation. During this session, the SLP played the role of a teacher and interacted with their own patient. In the control group, the SLP sat directly across from the participant and engaged in a face-to-face role-play. In the experimental group (*in virtuo*), the same SLP was "hidden" in a room adjacent to the experimental area and interacted with the participant in real time as a virtual avatar (Fig. 3). Thus, participants in the experimental group interacted with a virtual and unknown teacher, while those in the control group interacted with a real and familiar "teacher" (acted by their own SLP in a role-play context). In the control group, the SLP entered the room at the beginning of the conversation task and took a seat at the desk facing the participant. In the experimental group, the conversation task began when the virtual simulation was launched. To closely reflect the clinical reality in which such a tool might be utilized, participants were informed during the consent process that the virtual avatar would be controlled by their own SLP. However, this information was not emphasized further during the experiment.

In both conditions, the SLPs received identical instructions to ensure that the interaction remains as consistent and ecologically valid as possible from participant to participant. The scenario was: the SLP was a new teacher that just noticed that the participant was stuttering. The teacher listened to the participant's request and carried on discussions with them, initially interacting in a neutral and natural way. As the interaction progressed, anxiety-provoking stimuli (e.g., yawning, frowning, uncomfortable phrases) were introduced based on participant responses. The SLPs were given scripted statements to use during the interaction, such as "That's not fair to the other students" or "Why should I make special accommodations for you?". The SLPs followed the participant's behavior, not like systems where the stimuli are preset with scripts. They determined the most appropriate moment to introduce stimuli, considering the participant's responses and reactions, to keep the interaction as natural as possible. A member of the research team was directly observing the session in real time and marked the exact moment when each anxiety-inducing stimuli was presented, using the data acquisition software (AcqKnowledge).

Each SLP underwent a training session to provide them with recommendations such as waiting at least 20 s between the introduction of two stimuli (to facilitate accurate analysis of physiological responses) and not introducing two stimuli at the same time (to avoid confounding the data and to isolate the effect of each stimulus). They were also instructed on behaviors to avoid. For example, they were told not to discuss the differences between the types of scenarios being studied (e.g., by favoring or devaluing one condition over another), nor to reinforce reassurance-seeking or avoidance behaviors from the participant (e.g., avoid to say things like "speak more slowly,", or "you can switch the word if that one is too hard to say.").

2.4. Conversation task during session 2

The second session represents the *in vivo* condition with a room set up to simulate a classroom environment. The protocol for this session is similar to the first one, except that an actor assumes the role of the teacher instead of the SLP, so, the participant does not

Fig. 4. *In vivo* experimental condition. Participant seated at the desk and interacted with the real actor-teacher. The face of the participant has been masked to ensure anonymity.

know the "teacher". As in the first session for the control group, the actor sits directly across from the participant for a face-to-face interaction. The actor enters the room and takes his seat at the desk only at the beginning of the experimental condition (Fig. 4). During this second session, participants from both groups interact with a real and unknown actor-teacher.

3. Measures

3.1. Acceptability and presence (Objective 1)

Regarding objective 1, acceptability and perceived realism of the VR system were evaluated immediately after the *in virtuo* session using two complementary questionnaires: the **ITC-Sense of Presence Inventory (ITC-SOPI)** and the **Interaction with Avatars (IWA).**

The ITC-Sense of Presence Inventory (ITC-SOPI, Lessiter et al., 2001) is a 44-items self-report questionnaire designed to measure a user's subjective sense of presence during and after interaction with media or virtual environments (Baus et al., 2022; Lessiter et al., 2001). Responses are referring to the user's personal experience of perceiving the immersive experience as if it was real. The ITC-SOPI comprises four distinct subscales: Spatial Presence (assesses the extent to which users feel physically located within the virtual environment), Engagement (reflects psychological involvement, interest, and absorption in the experience), Ecological Validity (evaluates the perceived realism and fidelity of the virtual environment), and Negative Effects (measures symptoms of discomfort, such as disorientation or cybersickness). The ITC-SOPI has been established as a reliable and valid tool for assessing the sense of presence and shows a significant relationship with anxiety levels during virtual reality exposure therapy (Ling et al., 2014).

The Interaction with Avatars (IWA) questionnaire is a tailor-made tool specifically adapted for our study that assesses the users experience of interacting with a real-time animated avatar. The questionnaire consists of eight questions in total: five questions rated perceived realism of the virtual avatar interaction with a Likert scale from 0 to 10: 1. Did you feel like you were talking to a real person? 2. Did you feel like you were talking to your SLP? 3. Did the virtual simulation make you feel the same way as the real-life situation? 4. Were the practiced situations relevant to you? 5. Would you like to have access to this tool to practice in situations that you find stressful? The remaining three open-ended questions weighted the clinical utility of the user perspective: 6. What other situations would you like to work on using this tool? 7. What modifications could make the experience feel even closer to the real-life situation? 8. How many practice sessions with this tool would you like to do, or do you think would be necessary?

3.2. Physiological anxiety response (Objectives 2 and 3)

Physiological data were continuously recorded during each of the three phases (shaded in gray in Fig. 3): baseline, speech preparation and the conversational task. We utilized both **electrodermal activity (EDA)** and an **electrocardiogram (ECG)** for physiological indicators of anxiety.

3.2.1. Electodermal activity (EDA)

The use of EDA specifically adresses both our second and third objectives. EDA has been well-established as an indicator of anxiety when it comes to public speaking both our second and third objectives (Arsalan & Majid, 2021). It includes both **skin conductance level (SCL)** and the **skin conductance response (SCR)** to represent the activation of the sympathetic nervous system during stressors and ultimately is associated with stress levels in individuals (Kim et al., 2024).

SCL is the level of skin conductance recorded continuously across an interaction. It represents an overall level of arousal across time; higher SCL values mean more physiological arousal/stress levels. SCR is the rate of change of skin conductance that occurs in response to discrete stimuli, such as verbal or nonverbal communicative cues from the teacher. SCR is highly sensitive to short-term changes in emotions and can precisely track immediate anxiety responses to particular events.

In addition, we analyzed mean SCLs that occurred across conditions; thereby treated as a continuous measurement of arousal related to the speaking tasks. This approach allowed us to measure differences in arousal across conversational task and sessions, and addressed Objective 2 and 3, respectively.

3.2.2. Electrocardiogram (ECG)

The use of ECG specifically addresses our second and third objectives. ECG recordings complemented the EDA data by allowing calculation of **heart rate (HR)** and the Heart Rate Variability (HRV), represented by the **root mean square of successive differences between normal heartbeats (RMSSD)**. HR is the number of heart beats per minute (measured in bpm). It is the widely adopted measure to estimate levels of stress (Giannakakis et al., 2019). Heart Rate Variability (HRV) is a measure of time-based differences between adjacent heartbeats that reflects the heart's capacity to regulate and adapt to physiological and stress reactions. RMSSD is a calculated score to represent the short-term variability in heart rate and is used as an indicator of parasympathetic nervous system influence. It is considered the most common time-domain score to evaluate changes in HRV that are reflective of more parasympathetic nervous system impact (Shaffer & Ginsberg, 2017). Typically, decreased RMSSD and increased HR indicate elevated stress or anxiety (Michels et al., 2013).

3.2.3. Data acquisition

During all three conversational tasks (*in virtuo*, SLP role-play, and *in vivo*), the timing of each verbal and nonverbal stimulus (frowning, yawning and uncomfortable verbal remarks) introduced by the teacher was precisely annotated. The same member of the

research team, present during each session, annotated in real time the exact moment each stimulus was introduced using the Acq-Knowledge data acquisition software. In addition, video recordings of the sessions were reviewed afterward to verify that all stimulus events had been accurately recorded, and to identify any potential omissions. This double-checking process guaranteed that the data that were used for SCR analyses represented a consistent and comprehensive annotation of each anxiety-provoking cues. This detailed annotation process allowed us to isolate and analyze SCRs specifically related to each stimulus event and thus examine whether the virtual avatar could elicit physiological anxiety responses in participants, as described in Objective 2.

SCL and SCR were recorded using two EL507A Ag/AgCl electrodes filled with isotonic GEL 101A and connected to the Biopac MP36R system through SS57L cables. The electrodes were placed on the palmar surface of the index and middle fingers of the left hand, at the level of the first phalanges. Participants were instructed to remain still for 5–10 min before the start of the recordings to allow the gel to be absorbed and the signal to stabilize. For SCR acquisition, the minimum time separation between the stimulus event and the SCR peak was set to 1 s and the maximum to 6 s, in line with established guidelines (Braithwaite et al., 2013).

Heart rate and heart rate variability were recorded using three EL503 electrodes with GEL 100, connected through an SS2LB cable to the Biopac MP36R system. Because the study involved children and adolescents, we chose to avoid the standard clavicle–costal placement, which could be perceived as intrusive and induce stress among participants. Instead, an adapted peripheral configuration was used to minimize discomfort while maintaining signal quality: the positive (red) electrode was placed on the left leg, just above the ankle, and the negative (white) electrode on the right forearm, just above the wrist on the palmar side. The ground (black) electrode was not connected, as the reference signal was already provided through the electrodermal activity leads. It was taped under the table to keep it out of the participant's reach and to avoid any accidental contact with metallic surfaces.AcqKnowledge software (from Biopac) together with NeuroKit2 was used to extract and analyze ECG (HR, RMSSD) and EDA (SCL, SCR) data derived from Biopac recordings. NeuroKit2 is an open-source and user-centric Python library for neurophysiological signal processing (Makowski et al., 2021). The raw data were "cleaned" in AcqKnowledge to remove artifacts caused by electrode movements during the experiments. NeuroKit2's cleaned function was then used to process the EDA by removing noise and optimizing signal smoothing (using a low-pass filter with a 3 Hz cutoff frequency and a Butterworth filter), and to enhance ECG signal peak detection accuracy (with a 0.5 Hz high-pass Butterworth filter).Physiological measurements are not necessarily constant across days, or times in a day. For instance, a participant's resting SCL can be different across Session 1 and Session 2, due to natural changes in the physiological state. To account for this variability, these physiological measures (SCL, HR, and RMSSD) were all considered in terms of "change," relative to baseline recorded at the beginning of the session. In other words, the Session 1 data were related to a baseline from Session 1 and the Session 2 data were related to a baseline from Session 2. This tracks accounts for considerations as both within-session and between-session comparisons, as well as individual and contextual variation.

In contrast, SCR's are quick, short-term changes in skin conductance in response to a specific stimulus (e.g., frown or a verbal comment from the teacher). Unlike SCL, which reflects general arousal level, SCRs are event-related and occur relative to the participant's skin conductance level at the time of the stimulus. Because SCRs are defined as changes from this moment-to-moment baseline, they can be directly compared both within and across sessions, without requiring session-specific normalization.

3.3. Self-reported anxiety (Objectives 2 and 3)

To assess the participants' self-reported level of distress, we used the **Subjective Units of Distress Scale (SUDS).** The SUDS is a widely-adopted and valid measure of self-reported anxiety (Benjamin et al., 2010). The SUDS ranges from 0 to 100 with higher scores indicating greater anxiety. In this study, we used a modified SUDS (0−10), reflecting a rigorous and valid measure of anxiety in school-aged children (Bringuier et al., 2009).

To address the objectives 2 and 3, we asked the participants to complete the SUDS at three time points within each session (Fig. 3): (a) immediately after baseline, (b) immediately after speech preparation, before starting the conversation task and (c) immediately after the conversation task. For the baseline and speech preparation phases, participants were asked to rate, on a scale from 0 to 10, how they felt now. After the conversational task, participants were asked to retrospectively rate how they felt during the interaction with the teacher, at the beginning and at the end of the discussion, thus providing two separate SUDS ratings for this phase. This temporal profiling of SUDS ratings allows us to chart participants' perceived anxiety trajectories in each condition (*in virtuo* or SLP role-play, and *in vivo*). For the conversational task, we chose to administer the SUDS retrospectively rather than during the task, to avoid interfering with the task itself.

Additionally, we also collected SUDS ratings immediately pre- and post-placement of the physiological sensors so that we could control for any possible arousal shifts due to sensor placement and how it might have impacted baseline. Similarly, we obtained an additional SUDS rating after participants completed pre-experimental questionnaires (PRCS, LSAS-CA), which were completed after baseline measures to avoid any potential interference with baseline data. These supplementary ratings are not included in Fig. 3 to maintain clarity and focus on the main SUDS collection time points relevant to our primary objectives. Findings will be discussed in the Results section.

3.4. Self-efficacy (Objective 3)

To examine whether immersive VR practice enhances communicative confidence beyond conventional role-play, we used a **self-efficacy questionnaire**. In each session, participants rated their confidence before the speech preparation and after the conversational task. Differential pre–post scores reveal the extent to which avatar- or SLP-based practice yields gains in self-efficacy, directly testing Objective 3.

Public speaking self-efficacy was evaluated using a series of questions rated from 1 (“Not at all confident”) to 5 (“Very confident”), utilizing a self-efficacy questionnaire that was developed for this study. This questionnaire was adapted from (Bray et al., 2003), which was formerly utilized from (Manning, 1994) Self-Efficacy Scaling for Adolescents Who Stutter (SEA), adhering to (Bandura, 2006) framework.

The fourteen items on the questionnaire (see the Appendix) are divided into three categories. Section A is comprised of examples contextualized for the current study (e.g., *“I am confident talking about my stuttering in class with my teacher, even if they may have an unpleasant reaction”*), section B refers to school-based speaking situations (e.g., “I am confident asking a question in class”), and section C covers more general communication self-efficacy (e.g., “I am confident talking with a large group of my friends”).

We calculated a Cronbach's alpha coefficient of $\alpha = 0.87$, based on all the questions across all participants, indicating a high level of internal consistency among the items. This suggests that the questions are well correlated and effectively measure the same underlying concept.

3.5. Statistical analysis

Statistical analyses, described below, were conducted using IBM SPSS Statistics 29 and Python, with the Pandas, NumPy, SciPy, Matplotlib and Seaborn libraries, to examine both self-rated and quantitative variables across blocks and group configurations.

For SCRs, paired-sample *t*-tests were conducted to analyze the percentage of anxiety-provoking stimuli that elicited a skin conductance response in the experimental and control groups during both sessions 1 and 2. To analyze within- and between-group interactions and to increase the reliability of our statistical comparisons for physiological data (SCL, HR, and RMSSD), self-rated anxiety measures (SUDS) and self-efficacy, we used parametric and non-parametric bootstrap methods with 10,000 resamples. Bootstrap resampling provides robust estimates of the sampling distribution and mitigates problems associated with limited sample size, resulting in more accurate confidence intervals and significance tests (Hesterberg, 2014; Iwi et al., 1999). For continuous physiological measures (SCL, HR, RMSSD), a parametric bootstrap was conducted, with samples generated from a normal distribution with parameters estimated from the data. For questionnaire based variables (SUDS and self-efficacy), a non-parametric bootstrap was conducted, resampling directly from the observed data due to the variables discrete and ordinal nature.

Based on the bootstrap distributions obtained for each measure(SCL, SCR, HR, RMSSD, SUDS, and self-efficacy), we examined the mean and variability of the data for each variable at each phase of the experiment. Thus, we compared 95 % confidence intervals to examine within- and between-group differences for sessions 1 and 2. Non-overlapping 95 % confidence intervals were interpreted as evidence of statistically significant between-group differences, and represent a 5 % probability that the actual difference is outside the interval. Similarly, intervals that do not include the baseline value of 0 were considered to indicate significant within-group differences from baseline, at the 95 % confidence level. As applicable, p-values were calculated as the proportion of bootstrap resample where the statistic was as extreme or more extreme than the observed value. This approach enabled us to effectively demonstrate differences whether within-group and between-group, while reducing the number of tests we conducted and lowering the Type I error. To maintain clarity and brevity, in the Results section, we refer to bootstrap p-values, as simply p-values, unless stated. Given that bootstrap resampling provides empirical significance estimates without any parametric assumptions, we did not conduct any corrections for multiple comparisons (e.g. Bonferroni). Instead, we examined confidence intervals and bootstrap p-values, with emphasis on p-values near the threshold of significance.

In addition to significance tests, we have provided effect sizes (Cohen’s *d* and Pearson’s *r*) to quantify the magnitude of differences and relationships. We report effect size as absolute values to reflect the magnitude of the effect (e.g., small, medium, large effects), while the direction can be readily interpreted from the figures.

4. Results

4.1. Relevance and acceptability of the VR tool (Objective 1)

The results of the ITC-SOPI questionnaires reported a positive reception and substantial sense of presence in the participants ($n = 6$) during the virtual reality interaction with the virtual teacher. The “Spatial Presence” subscale indicates that participants experienced a good sense of being physically present in the virtual environment ($M = 3.60$, $SD = 0.70$). Engagement reflected a high level of psychological involvement and interest during the simulation ($M = 3.87$, $SD = 0.78$). The "Ecological Validity" subscale, which refers to

Table 2
Results from the five Likert-scale questions of the IWA questionnaire, with mean scores, standard deviations, and range for each item (Maximum possible score = 10).

IWA (questions 1–5)	*M*	*SD*	Range (min – max)
Did you feel like you were talking to a real person?	8.43	1.81	5–10
Did you feel like you were talking to your SLP?[a]	2.86	4.30	0–10
Did the VR simulation feel like a real situation?	6.50	3.45	0–10
Were the practiced situations relevant to you?	9.29	0.95	8–10
Would you like to have access to this tool to practice in situations you find stressful?	9.57	0.79	8–10

[a] Higher score on Item 2 reflects a reduced sense of presence, whereas for the other items, higher scores indicate higher sense of presence.

the perceived fidelity of the simulation, showed a good level of realism, but with larger variability between participants ($M = 3.27$, $SD = 1.53$). Interestingly, two participants rated this factor with the maximum score of 5, while another participant rated it as low as 1 out of 5. The same participant rated low scores for all four factors, suggesting a generally low engagement with the experience and also indicating a high degree of variability in perception between participants.

Finally, "Negative Effects", which includes potential disorientation, cybersickness or discomfort during the simulation, was rated as minimal by all participants ($M = 1.58$, $SD = 0.60$). This suggested that the VR simulation was well-tolerated and did not cause significant adverse effects.

The IWA questionnaire revealed several interesting results (see Table 2). First, participants indicated high levels of feeling that they are interacting with a real person, suggesting that the virtual environment was effective in simulating a real interaction. Second, the participants' perception of communicating with their own SLP was significantly low, suggesting that participants generally did not recognize or perceive that they were interacting with their own SLP while interacting with the avatar. This aspect is particularly interesting because it suggested that the setup (voice changes, facial expressions, and head movements) successfully allowed the SLP to embody the avatar without being recognized by the participants. However, given the high variability in responses, this result should be considered with caution. Although there was no significant correlation between scores on this question and participant age ($\rho = .281$, $p = .271$), the two highest scores (indicating recognition of their SLP) were given by the oldest participants (16 and 18 years old), with scores of 10 and 8, respectively.

These participants typically recognized their SLP's prosodic specifications (intonation and manner of speaking) rather than the voice itself. Third, there was a lot of variability in the participant's answers to the question asking if they felt like they were in a real situation (0, 1 participant, to 10, 2 participants) but participants perceived the simulated scenarios as highly relevant; indicating that the scenarios were appropriately matched to participant's needs. Finally, participants expressed a strong desire to access this VR tool for practicing stressful situations, highlighting the perceived usefulness of the VR environment for personal development and stress management.

Regarding the open-ended questions of the IWA questionnaire, the first one concerned other situations that participants would like to work with this VR tool. Some participants specifically expressed a desire to use this tool to practice oral presentations to a virtual class of students or to engage in conversation with a new person. Concerning the modifications that could make the experience feel closer to the real-life, two participants who were seen on the same day reported a noticeable lag between the voice and the avatar's lips, which was due to a weak Wi-Fi connection that particular day. Lastly, in regard to the number of practice sessions, participants wanted to use this tool much more, voicing a desire for once a day to twice practice prior to the actual *in vivo* presentation.

4.2. Skin conductance responses (Objective 2)

Table 3 presents the percentage of anxiety-provoking stimuli that elicited a skin conductance response, calculated as the number of stimuli followed by an SCR divided by the total number of stimuli presented. For example, during Session 1, participants in the experimental group reacted with SCRs to 20 % of the yawning stimuli introduced by the virtual teacher. Results are reported separately for each stimulus type (frowning, yawning, and disconcerting phrases), as well as for the overall total across all three categories.

Results suggested that the interactions with the virtual avatar can elicit physiological responses in participants. In the experimental group, no significant differences in the percentage of SCR detection were observed between the *in virtuo* and *in vivo* conditions ($t(4) = -1.07$, $p = .35$), indicating that the virtual avatar triggered responses comparable to those observed *in vivo* interactions.

During Session 1, frowning induced statistically more frequent SCRs in participants who talked face-to-face to their SLP (control group) compared to those who talked face-to-face with the virtual teacher (experimental group) ($t = -3.79$, $p < .05$).

Yawning and the disconcerting phrase, however, did not elicit significantly different SCR frequencies between the control and experimental groups ($t = -1.00$, $p = .347$; $t = -0.66$, $p = .526$, respectively). When considering all stimuli combined (see "Total" column in Table 3), participants in the control group exhibited significantly more frequent SCRs than those in the experimental group ($t = -3.76$, $p < .05$).

4.3. Self-reported anxiety (Objectives 2 and 3)

Additional SUDS ratings indicated that neither the placement of the physiological sensors nor the completion of the experimental questionnaires led to an increase in participants' arousal. This supports the use of baseline measures as valid reference points for assessing changes in self-reported anxiety throughout the experimental sessions.

SUDS scores during the speech preparation and conversational tasks were compared to the baseline score from the corresponding

Table 3
Means and standard deviations (SD) of the percentage of specific SCRs in both groups for Session 1 and Session 2.

		Anxiety-provoking stimuli eliciting a SCR			
	Group	% Frowning *M (SD)*	% Yawning *M (SD)*	% Disconcerting phrase *M (SD)*	% Total *M (SD)*
Session 1	Experimental	10.0 *(22.4)*	20.0 *(44.7)*	36.7 *(41.5)*	20.0 *(18.2)*
	Control	70.0 *(27.4)*	50.0 *(50.0)*	53.3 *(38.0)*	63.1 *(18.1)*
Session 2 *(in vivo)*	Experimental	25.0 *(50.0)*	25.0 *(50.0)*	60.0 *(54.7)*	35.0 *(35.6)*
	Control	70.0 *(44.7)*	50.0 *(57.7)*	50.0 *(50.0)*	60.0 *(36.5)*

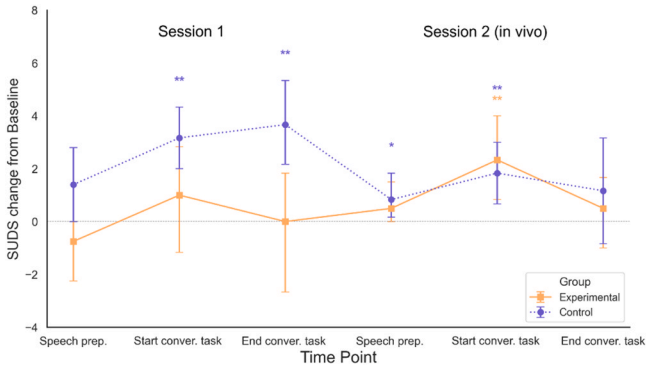


Fig. 5. Evolution of mean SUDS scores compared to the Baseline, across time points for control and experimental groups, for Session 1 (in virtuo or SLP) and Session 2 (*in vivo*). The dotted line at y = 0 represents the normalized baseline. Error bars represent 95 % bootstrap confidence intervals. Significant differences from baseline are indicated as: *p < .05, **p < .01 (bootstrap p-values).

session (Fig. 5). For Session 1 in the experimental group, conversation with the avatar (*in virtuo*) did not result in a significant change in self-rated stress from the baseline, at either the start (95 % CI = [-1.167, 2.833], $p = .40$), or end (95 % CI = [-2.667, 1.833], $p = 1.0$). In contrast, for the control group, conversation with the SLP had statistically significant higher self-rated stress than baseline, at start (95 % CI = [2.00, 4.33], $p < .001$) and at the end (95 % CI = [2.16, 5.16], $p < .001$) of the conversation, as indicated by non-overlapping 95 % confidence intervals Fig. 5. Additionally, during Session 2, at the start of the *in vivo* conversation, both the experimental (95 % CI = [0.83, 4.17], $p = .003$) and control groups (95 % CI = [0.67, 3.0], $p = .002$) showed significantly higher self-rated stress than baseline. By the end of the conversation, both groups had self-rated stress levels approaching baseline levels (experimental: 95 % CI = [−1.0, 1.67], $p = .56$; control: 95 % CI = [−0.83, 3.0], $p = .23$), suggesting a natural regulation of anxiety over the course of the interaction. This pattern will be discussed in Section 5.2.

When comparing the experimental and control groups, effect size analyses revealed large magnitude differences in SUDS ratings during Session 1 training ($d = 0.94$–1.35), with large effects decreasing to negligible differences during Session 2 *in vivo* exposure ($d = 0.25$–0.30). For Session 1, the SLP conversation at the end of the conversation was rated a higher stress level than the experimental group ($d = 1.35$, 95 % CI [–2.77, –0.67], $p = .031$), while other comparisons within Session 1 did demonstrate a large magnitude effect size (e.g., speech preparation, d = 1.27; conversation start, $d = 0.94$), despite p-values that did not reach significant differences. In Session 2, between-group effects were small ($d = 0.25$–0.30) and non-significant. These results highlighted that while statistical significance was limited, the magnitude of differences suggests that SLP conversations elicited stronger subjective stress responses than avatar interactions, especially during the Session 1.

4.4. Skin conductance levels (Objectives 2 and 3)

SCLs (specifically, continuous measures of skin conductance) were examined relative to each of the session baselines—that is, Session 1 data were compared to the baseline recorded at Session 1, and Session 2 data were compared to the baseline recorded at Session 2 (Fig. 6).

For the experimental group, the 95 % confidence intervals reflected that in session 1, both speech prep and conversation task with

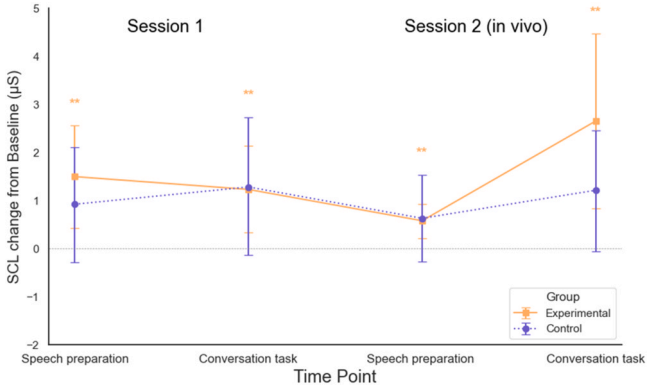


Fig. 6. Mean skin conductance level (SCL) deviation from baseline during speech preparation and conversation tasks by control and experimental groups, for Sessions 1 and 2. The dotted line at y = 0 represents the normalized baseline. Error bars represent 95 % bootstrap confidence intervals. Significant differences from baseline are indicated as: *p < .05, **p < .01 (bootstrap p-values).

virtual avatar significantly increased SCL from baseline level (95 % CI = [0.404, 2.578], $p = .006$ and 95 % CI = [0.351, 2.142], $p = .009$, respectively). The same pattern was seen in Session 2, where both the speech preparation phase and the conversation task with the actor-teacher also led to a significant increase in SCL (95 % CI = [0.215, 0.935], $p < .001$ and 95 % CI = [0.859, 4.471], $p = .008$, respectively). These increases were highlighted by a 95 % confidence interval that does not overlap the zero-baseline line, suggesting heightened arousal relative to initial activation measured at baseline.

In contrast, for the control group, the 95 % confidence intervals reflect that at both session 1 & 2, there was not a significant increase in speech preparation (95 % CI = [-0.267, 2.099], $p = .114$ and 95 % CI = [-0.280, 1.524], $p = .186$, respectively) or conversation task (with the SLP in Session 1: 95 % CI = [-0.156, 2.689], $p = .07$, and with the actor-teacher in Session 2: 95 % CI = [-0.041, 2.470], $p = .055$) from baseline.

Between-group effect size analyses indicated mostly negligible to small differences ($d = 0.06$–0.41) across most conditions, with the Session 2 conversational task showing a medium effect size ($d = 0.80$). Although the experimental group had significant ΔSCL relative to their own baselines across all conditions (all $p < .01$), these increases were not significantly different from the control group. These findings suggested that, despite within-group arousal changes from the baseline, between-group differences in SCL were generally minor.

4.5. Heart rate and heart rate variability (Objectives 2 and 3)

HR typically increases with increased stress, whereas RMSSD (which is calculated from heart rate variability) behaves inversely. In our study, Spearman analysis confirmed this strong inverse correlation between heart rate and RMSSD ($r = .683$, $p < .001$). For clarity of interpretation, we therefore chose to display inverse RMSSD (Fig. 8) to facilitate graphical interpretation in relation to arousal changes. Although both HR and RMSSD are continuous variables, the very small sample size (n = 6 per group) did not allow a reliable assessment of normality, and physiological measures such as HRV typically show non-normal distributions. We therefore used Spearman's rank correlation, which provides a robust estimate of monotonic associations under these conditions. Unexpectedly, for the experimental group, the 95 % confidence intervals indicated that the speech preparation and *in-vivo* conversation task with the avatar, in session 1, led to a significant decrease in HR (95 % CI = [-7.626, −0.755], $p = .010$ and 95 % CI = [-9.828, −2.325], $p = .002$, respectively) while significantly increasing RMSSD (95 % CI = [0.385, 7.275], $p = .027$ and 95 % CI = [4.699, 12.030], $p < .001$, respectively) as compared to baseline (Figs. 7 and 8). During Session 2, speech preparation also led to decreased HR (95 % CI = [-4.138, −0.944], $p = .001$), however, there was no difference observed in RMSSD (95 % CI = [-5.990, 3.955], $p = .709$) compared to baseline. During the *in vivo* conversation, both HR (95 % CI = [-4.157, 5.307], $p = .790$) and RMSSD (95 % CI = [-5.029, 5.565], $p = .898$) did not differ from baseline levels. There was no significant difference in HR and RMSSD between speech preparation and conversation task, for session 1 (HR: $p = .721$; RMSSD: $p = .310$) and for session 2 (HR: $p = .530$; RMSSD: $p = .893$).

Similarly, for the control group, during Session 1, both the speech preparation and the conversation task with the SLP led to lower HR (95 % CI = [-3.777, −0.699], $p = .002$ and 95 % CI = [-6.038, −2.029], $p < .001$, respectively), and an increased RMSSD compared to baseline for the conversational task (95 % CI = [1.139, 8.252], $p = .005$) but not for the speech preparation (95 % CI = [-0.926, 9.138], $p = .126$). During Session 2, the speech preparation induced a lower HR (95 % CI = [-4.736, −2.431], $p < .001$) and a higher RMSSD (95 % CI = [2.045, 18.248], $p = .013$) compared to baseline, also suggesting a reduction in arousal. These results are not what we predicted, as we planned for both conditions to increase physiological arousal. The *in vivo* conversation task, led to lower HR (95 % CI = [-6.877, −0.933], $p = .005$) but no significant difference in RMSSD (95 % CI = [-0.811, 11.886], $p = .098$) compared to baseline. Once again, no difference was noted in HR and RMSSD between speech preparation and conversation task for Session 1 (HR: $p = .602$; RMSSD: $p = .933$) and Session 2 (HR: $p = .963$; RMSSD: $p = .692$).

When comparing the experimental and control groups, effect size analyses revealed mostly medium to large differences in HR ($d = 0.55$–0.91) and negligeable to large differences in HRV ($d = 0.04$–1.33). For HRV during Session 2 speech preparation, a large and

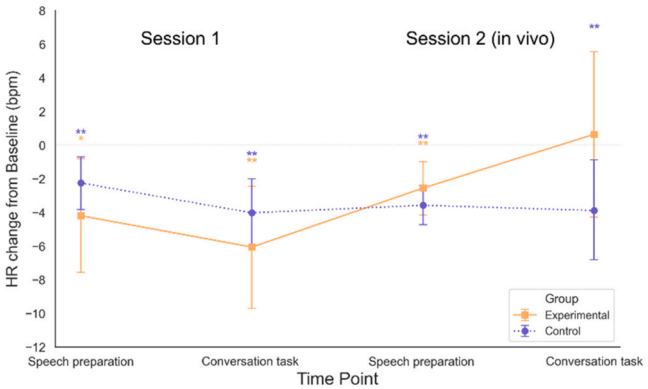


Fig. 7. Mean heart rate deviation from baseline during speech preparation and experimental conditions by control and experimental groups, for Sessions 1 and 2. The dotted line at $y = 0$ represents the normalized baseline. Error bars represent 95 % bootstrap confidence intervals. Significant differences from baseline are indicated as: $*p < .05$, $**p < .01$ (bootstrap p-values).

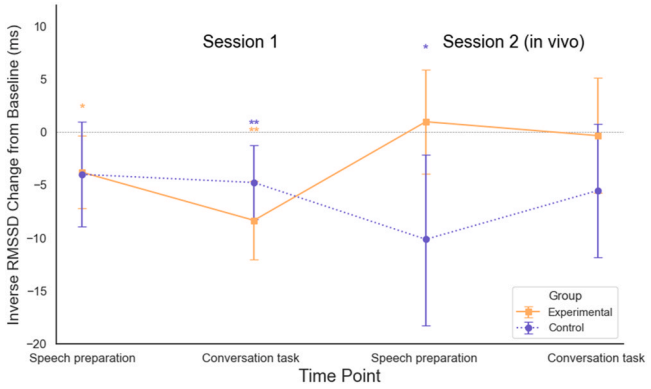


Fig. 8. Inverse RMSSD change from baseline during speech preparation and experimental conditions for control and experimental groups, for Sessions 1 and 2. The dotted line at $y = 0$ represents the normalized baseline. Error bars represent 95 % bootstrap confidence intervals. Significant differences from baseline are indicated as: $*p < .05$, $**p < .01$ (bootstrap p-values).

statistically significant effect was observed ($d = -1.33$, 95 % CI [-3.19, −0.19], $p = .045$), indicating higher RMSSD values in the control group compared to the experimental group. However, this finding should be interpreted with caution because the p-value is close to the significance level and the within-group confidence intervals shown in Fig. 8 overlap.

4.6. Self-efficacy (Objective 3)

Mann-Whitney U tests showed no significant differences in self-efficacy between the two groups across the Sessions, for any of the measurement times (i.e. before and after conversational task) in Session 1 ($U = 19.5$, $p = .873$, $r = 0.069$; $U = 20.0$, $p = .810$, $r = 0.092$, respectively), and before and after conversational task in Session 2 ($U = 21.5$, $p = .630$, $r = 0.162$; $U = 22.0$, $p = .589$, $r = 0.185$, respectively). These negligible to small effect sizes ($r = .069$–$.185$) indicate equivalent effectiveness between VR and role-play conditions (rather than insufficient statistical power).

The Wilcoxon signed-rank test indicated that the changes in self-efficacy within each group, between the initial measure (before conversation task in Session 1) and the last measure (after conversational task in Session 2) were not statistically significant for the experimental group ($Z = -1.36$, $p = .17$, $r = 0.60$) and for the control group ($Z = -0.94$, $p = .345$, $r = 0.39$). However, the moderate effect sizes ($r = 0.36$–0.60) for within-group changes suggested potential self-efficacy improvements that may require larger samples to achieve statistical significance.

Overall, the results indicate that neither training with the avatar nor with the SLP led to statistically significant increases in self-efficacy (Fig. 9). The moderate within-group effect sizes that were found suggested that both interventions might have therapeutic potential and should be studier further with larger samples. This conclusion holds true whether considering the total score based on the 14 items (rated on a 1-to-5 Likert scale) of the self-efficacy questionnaire or the three components (A, B, and C) assessed separately. These results will be discussed further in the Discussion section.

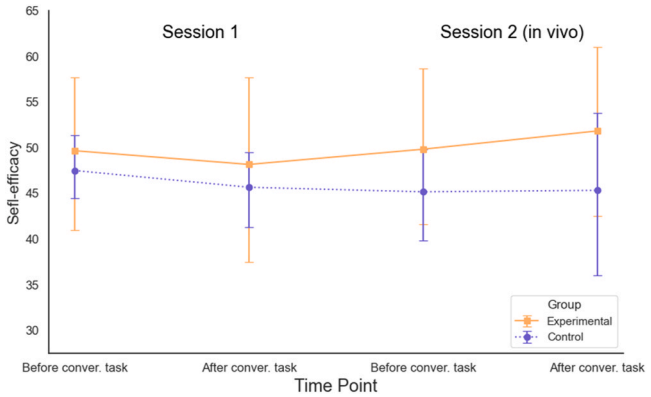


Fig. 9. Mean self-efficacy ratings with confidence intervals before and after conversation task for control and experimental groups across Sessions 1 and 2. The dotted line at $y = 0$ represents the normalized baseline. Error bars represent 95 % bootstrap confidence intervals.

5. Discussion

5.1. Objective 1: Acceptance and usability of the VR Tool

Our primary objective was to develop an immersive VR system designed to facilitate dynamic, real-time interactions between participants and a photorealistic virtual avatar. We assessed how well the system was accepted and perceived as relevant by children and adolescents who stutter. The findings from ITC-SOPI and IWA indicate good levels of acceptance, with high sense of presence, engagement, and low negative effects. The majority of the participants didn't recognize their own SLP during the conversation with the avatar, indicating the combination of voice modulation and avatar embodiment was effective. All participants reported that the simulated scenarios were very relevant to their needs and expressed enthusiasm about using the tool to practice difficult speaking situations, particularly with oral presentation.

Apart from the technical feasibility and acceptance of the system, these findings are important from the perspective of therapy. Acceptance and perceived realism are important conditions for VR to be a legitimate therapeutic support: if participants don't view the scenarios they are simulating as credible or engaging, then the ecological validity of the exposure is lessened, and therapeutic gain could be reduced (Parsons, 2015). This significance would suggest that the VR system might contribute to motivation and adherence with therapy.

Most participants rated perceived realism (relative to a real-life situation) positively, but one participant gave significantly low realism ratings. This variability among participants has been consistently noted in VR presence studies and may indicate differences in individual predisposition to become immersed in the virtual environment, that is the ability to feel present and engaged (Slater, 2003). Such differences are often linked to individual differences in cognitive style, imaginative engagement, past exposure to immersive technologies, or vulnerability to cybersickness (Witmer & Singer, 1998). Since our sample consisted of young participants, and in order to avoid adding to much cognitive load during the study, we did not use a more detailed questionnaire concerning presence, such as the Immersive Tendencies Questionnaire (Witmer & Singer, 1998). Future studies should consider incorporating these measures to more precisely analyse the influence of presence on the therapeutic effectiveness of VR in stuttering intervention programs.

5.2. Objective 2: Ability of the virtual teacher to elicit anxiety responses

Our second objective was to investigate whether the real-time, dynamic interactions with the virtual avatar could elicit physiological and self-reported anxiety responses in participants.

SCRs indicate that participants do not exhibit differences in anxiety-related responses between the *in virtuo* and *in vivo* conditions, which suggests that the anxiety-inducing stimuli were perceived similarly during the interaction with both the virtual and real teacher. Both the speech preparation and the conversation task also exhibited an increase in SCL as compared to baseline level which suggests increased arousal compared to the level of activation at baseline. Heart rate and heart rate variability (RMSSD) measures, however, suggested a more nuanced pattern, with reduction in arousal relative to baseline. This discrepancy highlights the dynamic nature of autonomic responses. HR may decrease due to adaptation or habituation, which is representative of the adaptive mechanisms of the autonomic nervous system during social stress (Kreibig, 2010).

While physiological data tend to show anxiety responses during conversations with the avatar, self-reported anxiety (SUDS) show no significant change compared to baseline. The anxiety that appears to be reflected in physiological parameters does not seem to be subjectively experienced by the participants. This divergence between physiological activation and self-reported anxiety is coherent with the framework of Lang's tripartite model of anxiety (Lang, 1968, 1979), delineates physiological arousal, cognitive (subjective) distress, and behavioral avoidance. Previous research, including studies with virtual audiences, has similarly shown that cognitive and physiological components of anxiety are not always closely linked and do not always change together in a predictable manner (Brundage et al., 2016; Girondini et al., 2023; Maron et al., 2016). In this study, several contextual factors may specifically contribute to this divergence. First, the relation between anxiety and the feeling of presence in virtual environments does not always appear to be straightforward, as highlighted by Felnhofer et al., (2014). Moreover, individuals with social anxiety may display biases in their perception of physiological arousal and social performance (Thomas & Johnco, 2025). Finally, emotion regulation and interoceptive abilities develop with age, and typically are limited in younger participants where the conscious ability to identify, differentiate, and verbalize anxiety may be limited. This mismatch has been documented in anxiety research where physiological arousal often gets misattributed, especially in children (Asbrand et al., 2020; Opdensteinen et al., 2021).

From a therapy or treatment perspective, this profile of increased physiological activation (SCL) without excessive subjective distress (SUDS) may actually be advantageous. Treatment engagement with social anxiety is often limited by the aversive quality of real-world exposure (Carl et al., 2019), thus, VR exposure may be a better initial starting point since it can produce the physiological mechanisms necessary for fear extinction, while lessening the perceived threat. This is especially important in a population of youth who are easily overwhelmed socially resulting in avoidant and disengaged responses.

5.3. Objective 3: Added value of VR as compared to traditional role-play

The third objective explored potential additional benefits of *in virtuo* practice compared to face-to-face role-play with an SLP.

First, regarding the comparison of experiences between the two training conditions, subjective anxiety ratings (SUDS) indicated that participants perceived virtual reality interactions as less anxiety inducing than real-life role-play with their SLP, before *in vivo* conversation. However, physiological measures presented a more nuanced picture. Skin conductance level (SCL) data suggested that

the physiological arousal of the experimental group from virtually interacting with the avatar was greater than the arousal of the control group with face-to-face talking with their own SLP. This suggests that the conversation with the virtual avatar triggered physiological arousal, but participants did not perceive this interaction as particularly anxiety-inducing. Similarly, skin conductance responses (SCRs) suggested some non-verbal stimuli, in particular frowning, were less perceived in the virtual teacher interaction than the SLP interaction. As frowning is a subtle non-verbal stimulus, it may be innately more difficult to perceive accurately from the virtual avatar than from a real person. Once again, heart rate (HR) and heart rate variability (RMSSD) showed reduced physiological arousal in both virtual and SLP interactions compared to baseline.

In terms of direct comparative effectiveness, specifically the value added by this VR environment compared to standard role-play with SLP, our results do not indicate an effect of a single training session (with either the virtual teacher or an SLP) on perceived self-efficacy or subsequent anxiety experienced when discussing with an unknown real actor-teacher. Although participants were interested in using the tool clinically, our study did not show any specific added benefits of VR compared to training with a SLP. However, caution is warranted in interpreting these results, given the small sample size, and limited number of training sessions.

Possible added benefits of using VR may not occur simply through reducing anxiety or increasing self-efficacy, but through adding the element of structured graduated exposure, within a structured therapeutic framework. VR can help children and adolescents to face increasingly difficult social speaking situations. By coupling VR practice with more traditional stuttering interventions (i.e. Cognitive Behavioral or Acceptance and Commitment Therapy), SLPs can use both the controlled, repeatable nature of VR and the personalized guidance of an SLP, to offer maximized skill/learning and emotional regulation.

This approach may improve treatment adherence through practicing skills within a safe, motivating, and engaging environment, potentially decreasing avoidance behaviors and enhancing generalization to real-world speaking situations. Future studies should explore longitudinal, multi-session interventions exploring VR and traditional therapy to see cumulative effects on self-efficacy, physiological arousal, and real-world communication performance.

6. Limitations and future directions

Despite the valuable results presented, there are several limitations to our pilot trial. First, due to the small sample size used in our study, the generalization of our findings may be restricted. There was a large amount of individual variation in the results: both perceived and experienced stress varied greatly among individuals. A larger sample would allow us to characterize this variability and better generalize or categorize profiles.

Another limitation is the number of training sessions: one training session was not enough to assess the impact of training on self-efficacy and anxiety. Longer intervention periods with more repeated exposures may have provided additional insight for those outcomes. Moreover, repeated training sessions would also allow for variability in training, which have impact on learning, including in VR (Lee et al., 2023; Nickel et al., 2019). The system could be used with different avatars, and in different scenarios, which further facilitates strong and decontextualized learning, to improve the transfer of skills to new contexts.

Another limitation involves the consistency of the actors. Due to availability of the staff, there were two different actors that played the classroom teacher in the study. Both actors were within the same age range and had similar body types. Importantly, both actors engaged with the participants in the control and experimental groups (they were not assigned to a specific group). However, even though actors used the same scenario and instructions, each actor naturally have different prosodic features and unique reactions. Moreover, each participant's own SLP, that they already knew, sometimes played the avatar and sometimes took part in the face-to-face role-play. Although varied interlocutors may have increased variability across participants, it is also representative of the natural variation found in a real clinical setting, where the SLP or teacher is rarely the same for every child.

Additionally, some participants recognized the prosody and speaking style of their own SLP, which may have influenced their emotional reactions. However, since everyone has their own intonation and speaking style, it makes it more difficult to manipulate those characteristics so that the participant does not recognize they are interfacing with their SLP. As mentioned in Section 4.12, while participants generally did not recognize or perceive that they were interacting with their own SLP while interacting with the avatar, the two oldest did report recognizing their SLP. Thus, future studies could further investigate the potential impact of participants' age on their engagement with the VR setup.

The SUDS was given after the conversation instead of during to keep the interaction uninterrupted, as previously noted in Moïse-Richard et al. (2021). However, this approach might present a constraint because participants with lower anxiety might have faced a possible distortion in retrospective self-reporting anxiety. They might have naturally regulated their anxiety during the task and, as a result, possibly underestimated the anxiety they initially experienced.

A further limitation is the absence of individual propensity for immersion measures. Differences in immersion propensity may account for variability in subjective presence and emotional responses, especially considering that immersion has been associated with a greater sense of presence, and that presence can influence emotional and behavioral responses in VR environments (Lønne et al., 2023). Future research would benefit from using validated questionnaires to measure individual immersion disposition, for example, the Immersive Tendencies Questionnaire (Witmer & Singer, 1998), to establish any potential impact it may have on emotion and physical responses in virtual environments. Although the original plan was to administer this questionnaire in the current study, the decision was made, following pilot testing, to exclude the questionnaire in order to reduce testing time and cognitive load, particularly for the youngest participants.

Some participants reported lower anxiety in one-on-one interactions than might be expected in group settings, suggesting that stress levels may vary with social context. Future studies could compare VR interactions with a single avatar to group interactions to examine the impact of social complexity on emotional response. Additionally, some children appeared to avoid eye contact with the

virtual teacher, which suggests that eye-tracking technology could be useful for measuring avoidance behaviors and improving comprehension of interaction dynamics.

Another factor to consider when interpreting our results is that the *in vivo* condition involved an unknown actor instead of participants' actual teacher. This choice of methodology helped maintain consistency for the participants, but it also diminished the personal stakes of the interaction. If a student declined their teacher's accommodation in a normal school setting, it would have real academic and social implications for them. In fact, they might experience reticence about declining, and the stigma of admitting they needed accommodations would have social implications. In the current study design, participant refusal had no real-life implications in the simulation. This may have, at least in part, contributed to lower perceived and self-reported levels of anxiety.

Finally, our study did not implement a full CBT exposure framework. The *in vivo* condition included an unknown actor instead of the participants' actual teacher, which improved experimental control but reduced personal stakes and might have contributed to lower perceived anxiety. The task mirrored public speaking tasks that produce anxiety symptoms similar to those used in exposure therapy but did not incorporate the entire therapeutic procedure used in CBT-based exposure. The procedure in this study would be better characterized as a "training session" rather than a formal exposure procedure. Future research with larger sample sizes, longer and repeated exposure, and incorporated full CBT procedures will be essential to assess the potential effectiveness of VR, both on its own or in combination with traditional exposure approaches, for anxiety reductions, self-efficacy, and real-life communication performance.

7. Conclusion

This pilot study investigated the feasibility and clinical relevance of an immersive virtual reality system that allowed children and adolescents who stutter to participate in real-time, dynamic conversation with a photorealistic virtual teacher. The system was well-received, promoted engagement and presence, and was considered by the participants to be relevant to "real world" communication. Results support the acceptance and ecological validity of the immersive VR system, as a potentially useful therapeutic intervention for speech and anxiety management.

Physiological data indicated that the virtual interaction elicited anxiety-related physical responses comparable to what could have been experienced in real-life social situations, although participants did not report significant increases in subjective anxiety. This discrepancy between physiological and self-report measures would, however, likely be helpful for treatment. VR exposure could activate fear extinction processes whilst minimizing the perception of threat that impedes participation in traditional exposure treatment.

This study did not demonstrate any additional value of VR training compared to role-play training with a SLP, at least when only one training session was offered. However, these findings support the potential of VR as a complementary tool available to SLPs, enabling precise control and adjustment of exposure parameters within CBT therapy, while also allowing clinicians to simulate repeated challenging speaking exercises in an immersive, engaging, and safe environment.

Future research should examine the impact of repeated exposure exercises using this virtual reality system, on perceived self-efficacy and anxiety reduction during subsequent real situations. Future work also should contain a broader range of speaking conditions, enhance the naturalness of the avatar and eliminate familiar SLP's prosody.

Although preliminary, this study suggests that immersive VR may represent an engaging and ecologically valid approach for practicing challenging speaking situations in youth who stutter, with potential to positively impact their everyday communication.

Declaration of Competing Interest

Stéphane Bouchard is consultant for, and holds equity in Cliniques et Développement In Virtuo, which develops virtual environments; however, Cliniques et Développement In Virtuo did not create the virtual environments used in this study. None of the authors have any conflicts of interest to declare.

Acknowledgements

We sincerely thank all the participants and their families who took part in this study. We are also grateful to the speech-language pathologists who contributed to the project, as well as to the teams from the Centre de réadaptation Marie Enfant of the CHU Sainte-Justine and the Centre de réadaptation en déficience physique Raymond-Dewar (IRD) for their collaboration. We would also like to thank Kevin Chenier for his initial contribution to the development of the tool. This work was supported by the Fonds de recherche du Québec (FRQ) through the AUDACE program (grant number 2022-AUDC-300126).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jfludis.2026.106194.

References

Andersen, N. J., Schwartzman, D., Martinez, C., Cormier, G., & Drapeau, M. (2023). Virtual reality interventions for the treatment of anxiety disorders: A scoping review. *Journal of Behavior Therapy and Experimental Psychiatry, 81*, Article 101851. https://doi.org/10.1016/j.jbtep.2023.101851

Anderson, P. L., Edwards, S. M., & Goodnight, J. R. (2016). Virtual reality and exposure group therapy for social anxiety disorder: Results from a 4–6 year follow-up. *Cognitive Therapy and Research, 41*(2), 230–236. https://doi.org/10.1007/s10608-016-9820-y

Anderson, P. L., Price, M., Edwards, S. M., Obasaju, M. A., Schmertz, S. K., Zimand, E., & Calamaras, M. R. (2013). Virtual reality exposure therapy for social anxiety disorder: A randomized controlled trial. *Journal of Consulting and Clinical Psychology, 81*(5), 751–760. https://doi.org/10.1037/a0033559

Arsalan, A., & Majid, M. (2021). Human stress classification during public speaking using physiological signals. *Computers in Biology and Medicine, 133*. https://doi.org/10.1016/j.compbiomed.2021.104377

Asbrand, J., Schulz, A., Heinrichs, N., & Tuschen-Caffier, B. (2020). Biased perception of physiological arousal in child social anxiety disorder before and after cognitive behavioral treatment. *Clinical Psychology in Europe, 2*. https://doi.org/10.32872/cpe.v2i2.2691

Bandura, A. (1986). Social Foundations of Thought & Action: A Social Cognitive Theory. In Prentice-Hall series in social learning theory.

Bandura, A. (2006). Guide for constructing self-efficacy scales (revised). *Self-Efficacy Beliefs of Adolescents, 5*, 307–337.

Baus, O., Bouchard, S., Nolet, K., & Berthiaume, M. (2022). In a dirty virtual room: Exposure to an unpleasant odor increases the senses of presence, reality, and realism. *Cogent Psychology, 9*(1), 2115690. https://doi.org/10.1080/23311908.2022.2115690

Baxter, S., Johnson, M., Blank, L., Cantrell, A., Brumfitt, S., Enderby, P., & Goyder, E. (2016). Non-pharmacological treatments for stuttering in children and adults: A systematic review and evaluation of clinical effectiveness, and exploration of barriers to successful outcomes. *Health Technology Assessment, 20*, 1–83. https://doi.org/10.3310/hta20020. xi.

Beele, G., Liesong, P., Bojanowski, S., Hildebrand, K., Weingart, M., Asbrand, J., Correll, C. U., Morina, N., & Uhlhaas, P. J. (2024). Virtual reality exposure therapy for reducing school anxiety in adolescents: Pilot study. *JMIR Ment Health, 11*, Article e56235. https://doi.org/10.2196/56235

Bell, I., Nicholas, J., Alvarez-Jimenez, M., Thompson, A., & Valmaggia, L. (2020). Virtual reality as a clinical tool in mental health research and practice. *Dialogues in Clinical Neuroscience, 22*, 169–177. https://doi.org/10.31887/DCNS.2020.22.2/lvalmaggia

Benjamin, C., Wolk, O'Neil, K., Crawley, S., Beidas, R., Coles, M., & Kendall, P. (2010). Patterns and predictors of subjective units of distress in anxious youth. *Behavioural and Cognitive Psychotherapy, 38*, 497–504. https://doi.org/10.1017/S1352465810000287

Bernard, R., Hofslundsengen, H., & Norbury, C. (2022). Anxiety and depression symptoms in children and adolescents who stutter: A systematic review and meta-analysis. *Journal of Speech, Language, and Hearing Research, 65*, 1–21. https://doi.org/10.1044/2021_JSLHR-21-00236

Blood, G., Blood, I., Tramontana, G., Sylvia, A., Boyle, M., & Motzko, G. (2011). Self-reported experience of bullying of students who stutter: Relations with life satisfaction, life orientation, and self-esteem. *Perceptual and Motor Skills, 113*, 353–364. https://doi.org/10.2466/07.10.15.17.PMS.113.5.353-364

Bouchard, S., Côté, S., & Richard, D. C. S. (2007). Chapter 16—Virtual reality applications for exposure. In D. C. S. Richard, & D. Lauterbach (Eds.), *Handbook of Exposure Therapies* (pp. 347–388). Academic Press. https://doi.org/10.1016/B978-012587421-2/50017-X.

Bouchard, S., Dumoulin, S., Robillard, G., Guitard, T., Klinger, É., Forget, H., Loranger, C., & Roucaut, F. X. (2017). Virtual reality compared with *in vivo* exposure in the treatment of social anxiety disorder: A three-arm randomised controlled trial. *The British Journal of Psychiatry, 210*(4), 276–283. https://doi.org/10.1192/bjp.bp.116.184234

Bouchard, S., Mendlowitz, S. L., Coles, M. E., & Franklin, M. (2004). Considerations in the use of exposure with children. *Cognitive and Behavioral Practice, 11*(1), 56–65. https://doi.org/10.1016/S1077-7229(04)80007-5

Braithwaite, J., Watson, D., Jones, R., & Rowe, M.A. (2013). Guide for Analysing Electrodermal Activity & Skin Conductance Responses for Psychological Experiments. CTIT technical reports series.

Bray, M., Kehle, T., Lawless, K., & Theodore, L. (2003). The relationship of self-efficacy and depression to stuttering. *American Journal of Speech-Language Pathology / American Speech-Language-Hearing Association, 12*, 425–431. https://doi.org/10.1044/1058-0360(2003/088)

Bringuier, S., Dadure, C., Raux, O., Dubois, A., Picot, M.-C., & Capdevila, X. (2009). The perioperative validity of the visual analog anxiety scale in children: A discriminant and useful instrument in routine clinical practice to optimize postoperative pain management. *Anesthesia Analgesia, 109*(3). ⟨https://journals.lww.com/anesthesia-analgesia/fulltext/2009/09000/the_perioperative_validity_of_the_visual_analog.9.aspx⟩.

Brundage, S., Brinton, J., & Hancock, A. (2016). Utility of virtual reality environments to examine physiological reactivity and subjective distress in adults who stutter. *Journal of Fluency Disorders, 50*, 85–95. https://doi.org/10.1016/j.jfludis.2016.10.001

Brundage, S., & Graap, K. (2004). Virtual reality: An exciting new tool to enhance stuttering treatment. *Perspectives on Fluency and Fluency Disorders, 14*, 4–9. https://doi.org/10.1044/ffd14.2.4

Brundage, S., Graap, K., Gibbons, K., Ferrer, M., & Brooks, J. (2006). Frequency of stuttering during challenging and supportive virtual reality job interviews. *Journal of Fluency Disorders, 31*, 325–339. https://doi.org/10.1016/j.jfludis.2006.08.003

Brundage, S., & Hancock, A. (2015). Real enough: Using virtual public speaking environments to evoke feelings and behaviors targeted in stuttering assessment and treatment. *American Journal of Speech-Language Pathology / American Speech-Language-Hearing Association, 24*. https://doi.org/10.1044/2014_AJSLP-14-0087

Carl, E., Stein, A. T., Levihn-Coon, A., Pogue, J. R., Rothbaum, B., Emmelkamp, P., Asmundson, G. J. G., Carlbring, P., & Powers, M. B. (2019). Virtual reality exposure therapy for anxiety and related disorders: A meta-analysis of randomized controlled trials. *Journal of Anxiety Disorders, 61*, 27–36. https://doi.org/10.1016/j.janxdis.2018.08.003

Cecil, J., Tetnowski, J., & Tentu, S.K. (2024). A Virtual Reality based Therapeutic Approach for Stuttering Intervention.

Chambers, C., & Johnston, C. (2002). Developmental differences in children's use of rating scales. *Journal of pediatric Psychology, 27*, 27–36. https://doi.org/10.1093/jpepsy/27.1.27

Chard, I., Van Zalk, N., & Picinali, L. (2023). Virtual reality exposure therapy for reducing social anxiety associated with stuttering: The role of outcome expectancy, therapeutic alliance, presence and social presence. *Frontiers in Virtual Reality, 4*. https://doi.org/10.3389/frvir.2023.1159549

Chard, I., van Zalk, N., & Picinali, L. (2023). Virtual reality exposure therapy for reducing social anxiety in stuttering: A randomized controlled pilot trial. *Frontiers in Digital Health, 5*. https://doi.org/10.3389/fdgth.2023.1061323

Chesham, R. K., Malouff, J. M., & Schutte, N. S. (2018). Meta-analysis of the efficacy of virtual reality exposure therapy for social anxiety. *Behaviour Change, 35*(3), 152–166. https://doi.org/10.1017/bec.2018.15

Chittaro, L., & Serafini, M. (2023). Desktop virtual reality as an exposure method for test anxiety: Quantitative and qualitative feasibility study. *Multimedia Tools and Applications*, 1–27. https://doi.org/10.1007/s11042-023-16917-2

Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. *Social phobia: Diagnosis, Assessment, and Treatment* (pp. 69–93).

Côté, S., & Bouchard, S. (2009). Cognitive mechanisms underlying virtual reality exposure. *Cyberpsychology Behavior: the Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society, 12*, 121–129. https://doi.org/10.1089/cpb.2008.0008

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy, 58*, 10–23. https://doi.org/10.1016/j.brat.2014.04.006

Dadds, M., Perrin, S., & Yule, W. (1998). Social desirability and self-reported anxiety in children: An analysis of the RCMAS lie scale. *Journal of Abnormal Child Psychology, 26*, 311–317. https://doi.org/10.1023/A:1022610702439

El-Yamri, M., Romero-Hernandez, A., Gonzalez-Riojo, M., & Manero, B. (2019). Designing a VR game for public speaking based on speakers features: A case study. *Smart Learning Environments, 6*(1), 12. https://doi.org/10.1186/s40561-019-0094-1

Erickson, S., & Block, S. (2013). The social and communication impact of stuttering on adolescents and their families. *Journal of Fluency Disorders, 38*(4), 311–324. https://doi.org/10.1016/j.jfludis.2013.09.003

Etienne, E., Leclercq, A.-L., Remacle, A., Dessart, L., & Schyns, M. (2023). Perception of avatars nonverbal behaviors in virtual reality. *Psychology Marketing, 40*. https://doi.org/10.1002/mar.21871

Felnhofer, A., Kothgassner, O. D., Hetterle, T., Beutl, L., Hlavacs, H., & Kryspin-Exner, I. (2014). Afraid to be there ? Evaluating the relation between presence, self-reported anxiety, and heart rate in a virtual public speaking task. *Cyberpsychology, Behavior, and Social Networking, 17*(5), 310–316. https://doi.org/10.1089/cyber.2013.0472

Gerlach-Houck, H., Kubart, K., & Cage, E. (2022). Concealing stuttering at school: "When you can't fix it…the only alternative is to hide it". *Language, Speech, and Hearing Services in Schools, 54*, 1–18. https://doi.org/10.1044/2022_LSHSS-22-00029

Giannakakis, G., Grigoriadis, D., Giannakaki, K., Simantiraki, O., Roniotis, A., & Tsiknakis, M. (2019). Review on psychological stress detection using biosignals. *IEEE Transactions on Affective Computing* (p. 1). https://doi.org/10.1109/TAFFC.2019.2927337

Girondini, M., Frigione, I., Marra, M., Stefanova, M., Pillan, M., Maravita, A., & Gallace, A. (2024). Decoupling the role of verbal and non-verbal audience behavior on public speaking anxiety in virtual reality using behavioral and psychological measures. *Frontiers in Virtual Reality*, *5*. https://doi.org/10.3389/frvir.2024.1347102

Girondini, M., Stefanova, M., Pillan, M., & Gallace, A. (2023). Speaking in front of cartoon avatars: A behavioral and psychophysiological study on how audience design impacts on public speaking anxiety in virtual environments. *International Journal of Human-Computer Studies, 179*, Article 103106. https://doi.org/10.1016/j.ijhcs.2023.103106

Glémarec, Y. (2023). Audience simulation and perception in virtual reality.: Application to public speaking training in an academic environment. École Nationale d'Ingénieurs de Brest.

Glémarec, Y., Lugrin, J.-L., Bosser, A.-G., Collins Jackson, A., Buche, C., & Latoschik, M. E. (2021). Indifferent or enthusiastic ? Virtual audiences animation and perception in virtual reality. *Frontiers in Virtual Reality*, *2*. https://doi.org/10.3389/frvir.2021.666232

Heeren, A., Ceschi, G., Valentiner, D., Dethier, V., & Philippot, P. (2013). Assessing public speaking fear with the short form of the personal report of confidence as a speaker scale: Confirmatory factor analyses among a French-speaking community sample. *Neuropsychiatric Disease and Treatment, 9*, 609–618. https://doi.org/10.2147/NDT.S43097

Hesterberg, T. (2014). What teachers should know about the bootstrap: Resampling in the undergraduate statistics curriculum. *The American Statistician*, *69*. https://doi.org/10.1080/00031305.2015.1089789

Horigome, T., Kurokawa, S., Sawada, K., Kudo, S., Shiga, K., Mimura, M., & Kishimoto, T. (2020). Virtual reality exposure therapy for social anxiety disorder: A systematic review and meta-analysis. *Psychological Medicine, 50*(15), 2487–2497. https://doi.org/10.1017/S0033291720003785

Iverach, L., Jones, M., McLellan, L. F., Lyneham, H. J., Menzies, R. G., Onslow, M., & Rapee, R. M. (2016). Prevalence of anxiety disorders among children who stutter. *Journal of Fluency Disorders, 49*, 13–28. https://doi.org/10.1016/j.jfludis.2016.07.002

Iverach, L., & Rapee, R. M. (2014). Social anxiety disorder and stuttering: Current status and future directions. *Journal of Fluency Disorders, 40*, 69–82. https://doi.org/10.1016/J.JFLUDIS.2013.08.003

Iwi, G., Millard, R., Palmer, A., Preece, A., & Saunders, M. (1999). Bootstrap resampling: A powerful method of assessing confidence intervals for doses from experimental data. *Physics in Medicine and Biology, 44*, N55–N62. https://doi.org/10.1088/0031-9155/44/4/021

Johnson, C., Gerwin, K., Tichenor, S., Boyle, M., & Walsh, B. (2024). Evaluating stuttering self-stigma and its relationship to adverse impact in children and adolescents with the child stuttering self-stigma scale. *Journal of Speech, Language, and Hearing Research, 67*, 2920–2934. https://doi.org/10.1044/2024_JSLHR-24-00069

Johnson, G., Onslow, M., Horton, S., & Kefalianos, E. (2023b). Reduced stuttering for school-age children: A systematic review. *Journal of Fluency Disorders, 78*, Article 106015. https://doi.org/10.1016/j.jfludis.2023.106015

Johnson, G., Onslow, M., Horton, S., & Kefalianos, E. (2023a). Psychosocial features of stuttering for school-age children: A systematic review. *International Journal of Language Communication Disorders, 58*. https://doi.org/10.1111/1460-6984.12887

Jones, M., Menzies, R., Onslow, M., Lowe, R., O'Brian, S., & Packman, A. (2021). Measures of psychological impacts of stuttering in young school-age children: A systematic review. *Journal of Speech, Language, and Hearing Research, 64*, 1918–1928. https://doi.org/10.1044/2021_JSLHR-20-00455

Kampmann, I. L., Emmelkamp, P. M. G., Hartanto, D., Brinkman, W.-P., Zijlstra, B. J. H., & Morina, N. (2016). Exposure to virtual social interactions in the treatment of social anxiety disorder: A randomized controlled trial. *Behaviour Research and Therapy, 77*, 147–156. https://doi.org/10.1016/j.brat.2015.12.016

Kelly, O., Matheson, K., Martinez, A., Merali, Z., & Anisman, H. (2007). Psychosocial stress evoked by a virtual audience: Relation to neuroendocrine activity. *Cyberpsychology Behavior: the Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society, 10*, 655–662. https://doi.org/10.1089/cpb.2007.9973

Kim, D., Kim, H., Kim, K., Kim, M., & Jeon, H. J. (2024). P51: Correlation between skin conductance and anxiety in virtual reality. *International Psychogeriatrics, 35*, 154–155. https://doi.org/10.1017/S1041610223002910

Kiser, D., Gromer, D., Pauli, P., & Hilger, K. (2022). A virtual reality social conditioned place preference paradigm for humans: Does trait social anxiety affect approach and avoidance of virtual agents? *Frontiers in Virtual Reality, 3*, Article 916575. https://doi.org/10.3389/frvir.2022.916575

Kohmäscher, A., Primaßin, A., Heiler, S., Avelar, P., Franken, M.-C., & Heim, S. (2023). Effectiveness of stuttering modification treatment in school-age children who stutter: A randomized clinical trial. *Journal of Speech, Language, and Hearing Research, 66*, 4191–4205. https://doi.org/10.1044/2023_JSLHR-23-00224

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion : A review. *Biological Psychology, 84*(3), 394–421. https://doi.org/10.1016/j.biopsycho.2010.03.010

Krumhuber, E., Skora, L., Hill, H., & Lander, K. (2023). The role of facial movements in emotion recognition. *Nature Reviews Psychology*, *2*. https://doi.org/10.1038/s44159-023-00172-1

Lang, P. J. (1968). Fear reduction and fear behavior: Problems in treating a construct. *Research in psychotherapy*, 90–102. https://doi.org/10.1037/10546-004

Lang, P. J. (1979). A bio-informational theory of emotional imagery. *Psychophysiology, 16*(6), 495–512. https://doi.org/10.1111/j.1469-8986.1979.tb01511.x

Lee, Y. S., Rashidi, A., Talei, A., Jian, B., & Rashidi, S. (2023). A Comparison Study on the Learning Effectiveness of Construction Training Scenarios in a Virtual Reality Environment. *Virtual Worlds, 2*. https://doi.org/10.3390/virtualworlds2010003

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-sense of presence inventory. *Presence, 10*, 282–297. https://doi.org/10.1162/105474601300343612

Leyns, C., Bosschem, L., Papeleu, T., Sabbe, L., Walkom, G., & D'haeseleer, E. (2025). Virtual reality as a tool in gender-affirming voice training: A pilot study. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2025.06.034

Li, M., Pan, J., Gao, Y., Shen, Y., Luo, F., Dai, J., Hao, A., & Qin, H. (2022). Neurophysiological and subjective analysis of VR emotion induction paradigm. *IEEE Transactions on Visualization and Computer Graphics* (pp. 1–11). https://doi.org/10.1109/TVCG.2022.3203099

Ling, Y., Nefs, H., Morina, N., Heynderickx, I., & Brinkman, W.-P. (2014). A meta-analysis on the relationship between self-reported presence and anxiety in virtual reality exposure therapy for anxiety disorders. *PloS One, 9*, Article e96144. https://doi.org/10.1371/journal.pone.0096144

Lønne, T. F., Karlsen, H. R., Langvik, E., & Saksvik-Lehouillier, I. (2023). The effect of immersion on sense of presence and affect when experiencing an educational scenario in virtual reality: A randomized controlled study. *Heliyon, 9*(6), Article e17196. https://doi.org/10.1016/j.heliyon.2023.e17196

Makowski, D., Pham, T., Lau, Z. J., Brammer, J. C., Lespinasse, F., Pham, H., Schölzel, C., & Chen, S. H. A. (2021). NeuroKit2: A Python toolbox for neurophysiological signal processing. *Behavior Research Methods, 53*(4), 1689–1696. https://doi.org/10.3758/s13428-020-01516-y

Man, S. S., Xiaoyu, L., Xiao Jing, L., Yu-Chi, L., & Chan, A. H. S. (2024). Assessing the effectiveness of virtual reality interventions on anxiety, stress, and negative emotions in college students: A meta-analysis of randomized controlled trials. *International Journal of Human–Computer Interaction*, 1–17. https://doi.org/10.1080/10447318.2024.2434957

Manning, W. (1994). The SEA-Scale: Self-efficacy scaling for adolescents who stutter. annual meeting of the American Speech-Language-Hearing Association, New Orleans, LA.

Maron, P., Powell, V., & Powell, W. (2016). The differential effect of neutral and fear-stimulus virtual reality exposure on physiological indicators of anxiety in acrophobia.

Masia Warner, C., Storch, E., Pincus, D., Klein, R., Heimberg, R., & Liebowitz, M. (2003). The Liebowitz Social Anxiety Scale for Children and Adolescents: An Initial Psychometric Investigation. *Journal of the American Academy of Child and Adolescent Psychiatry, 42*, 1076–1084. https://doi.org/10.1097/01.CHI.0000070249.24125.89

Menzies, R. G., Onslow, M., Packman, A., & O'Brian, S. (2009). Cognitive behavior therapy for adults who stutter: A tutorial for speech-language pathologists. *Journal of Fluency Disorders, 34*(3), 187–200. https://doi.org/10.1016/j.jfludis.2009.09.002
Michels, N., Sioen, I., Clays, E., De Buyzere, M., Ahrens, W., Huybrechts, I., Vanaelst, B., & Henauw, S. (2013). Children's heart rate variability as stress indicator: Association with reported stress and cortisol. *Biological Psychology*, *94*. https://doi.org/10.1016/j.biopsycho.2013.08.005
Miers, A., Blöte, A., Sumter, S., Kallen, V., & Westenberg, P. M. (2011). Subjective and objective arousal correspondence and the role of self-monitoring processes in high and low socially anxious youth. *Journal of Experimental Psychopathology, 2*, 531–550. https://doi.org/10.5127/jep.019411
Miluczenko, H., Dohmen, A., St., Louis, K., & Sommer, S. (2025). German non-stuttering primary school children's knowledge and attitudes towards stuttering: International comparison and influencing factors. *International Journal of Language Communication Disorders*, *60*. https://doi.org/10.1111/1460-6984.70083
Moïse-Richard, A., Ménard, L., Bouchard, S., & Leclercq, A.-L. (2021). Real and virtual classrooms can trigger the same levels of stuttering severity ratings and anxiety in school-age children and adolescents who stutter. *Journal of Fluency Disorders, 68*, Article 105830. https://doi.org/10.1016/j.jfludis.2021.105830
Murphy, W. P., Yaruss, J. S., & Quesal, R. W. (2007). Enhancing treatment for school-age children who stutter: II. Reducing bullying through role-playing and self-disclosure. *Journal of Fluency Disorders, 32*(2), 139–162. https://doi.org/10.1016/j.jfludis.2007.02.001
Nickel, C., Knight, C., Langille, A., & Godwin, A. (2019). How Much Practice Is Required to Reduce Performance Variability in a Virtual Reality Mining Simulator? *Safety, 5*(2), 18. https://doi.org/10.3390/safety5020018
Opdensteinen, K. D., Schaan, L., Pohl, A., Schulz, A., Domes, G., & Hechler, T. (2021). Interoception in preschoolers: New insights into its assessment and relations to emotion regulation and stress. *Biological Psychology, 165*, Article 108166. https://doi.org/10.1016/j.biopsycho.2021.108166
Ornstein, A. F., & Manning, W. H. (1985). Self-efficacy scaling by adult stutterers. *Journal of Communication Disorders, 18*(4), 313–320. https://doi.org/10.1016/0021-9924(85)90008-5
Pan, X., & Hamilton, A. F. de C. (2018). Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. *British Journal of Psychology, 109*(3), 395–417. https://doi.org/10.1111/bjop.12290
Parrish, D. E., Oxhandler, H. K., Duron, J. F., Swank, P., & Bordnick, P. (2016). Feasibility of virtual reality environments for adolescent social anxiety disorder. *Research on Social Work Practice, 26*(7), 825–835. https://doi.org/10.1177/1049731514568897
Parsons, T. D. (2015). Virtual reality for enhanced ecological validity and experimental control in the clinical, affective and social neurosciences. *Frontiers in Human Neuroscience, 9-2015*. https://doi.org/10.3389/fnhum.2015.00660
Premkumar, P., Heym, N., Brown, D., Battersby, S., Sumich, A., Huntington, B., Daly, R., & Zysk, E. (2021). The effectiveness of self-guided virtual-reality exposure therapy for public-speaking anxiety. *Frontiers in Psychiatry*, *12*. https://doi.org/10.3389/fpsyt.2021.694610
Rapee, R. M., & Heimberg, R. G. (1997). A cognitive-behavioral model of anxiety in social phobia. *Behaviour Research and Therapy, 35*(8), 741–756. https://doi.org/10.1016/S0005-7967(97)00022-3
Schmits, E., Heeren, A., & Quertemont, E. (2014). The self-report version of the LSAS-CA: Psychometric properties of the French version in a non-clinical adolescent sample. *Psychologica Belgica, 54*, 181–198. https://doi.org/10.5334/pb.al
Schöne, B., Kisker, J., Lange, L., Gruber, T., Sylvester, S., & Osinsky, R. (2023). The reality of virtual reality. *Frontiers in Psychology, 14*. https://doi.org/10.3389/fpsyg.2023.1093014
Seymour, M., Yuan, L., Dennis, A., & Riemer, K. (2021). Have we crossed the Uncanny Valley ? Understanding affinity, trustworthiness, and preference for realistic digital humans in immersive environments. *Journal of the Association for Information Systems, 22*, 591–617. https://doi.org/10.17705/1jais.00674
Shaffer, F., & Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. *Frontiers in Public Health, 5*. https://doi.org/10.3389/fpubh.2017.00258
Slater, M. (2003). A note on presence terminology. *Presence Connect, 3*.
Smith, K. A., Iverach, L., O'Brian, S., Kefalianos, E., & Reilly, S. (2014). Anxiety of children and adolescents who stutter: A review. *Journal of Fluency Disorders, 40*, 22–34. https://doi.org/10.1016/j.jfludis.2014.01.003
Son, G., & Rubo, M. (2025). Social virtual reality elicits natural interaction behavior with self-similar and generic avatars. *International Journal of Human-Computer Studies, 199*, Article 103488. https://doi.org/10.1016/j.ijhcs.2025.103488
Sprecher, S., & Hampton, A. J. (2017). Liking and other reactions after a get-acquainted interaction: A comparison of continuous face-to-face interaction versus interaction that progresses from text messages to face-to-face. *Communication Quarterly, 65*(3), 333–353. https://doi.org/10.1080/01463373.2016.1256334
Stangier, U., Heidenreich, T., Peitz, M., Lauterbach, W., & Clark, D. (2003). Cognitive therapy for social phobia: Individual versus group treatment. *Behaviour Research and Therapy, 41*, 991–1007. https://doi.org/10.1016/S0005-7967(02)00176-6
Tentu, S.K., & Cecil, J. (2024). The Potential of virtual reality Digital Twins to serve as therapy approaches for stuttering (p. 9). https://doi.org/10.1109/SeGAH61285.2024.10639574.
Thomas, T., & Johnco, C. (2025). Biased perceptions of physiological arousal in social anxiety: Understanding the role of objective and subjective physiological arousal in the discrepancy between self and observer perceptions of social performance. *Cognitive Therapy and Research, 49*, 824–834. https://doi.org/10.1007/s10608-025-10583-4
Trautmann, S. A., Fehr, T., & Herrmann, M. (2009). Emotions in motion: Dynamic compared to static facial expressions of disgust and happiness reveal more widespread emotion-specific activations. *Brain Research, 1284*, 100–115. https://doi.org/10.1016/j.brainres.2009.05.075
Tse, Z., Emad, S., Hasan, M., Papathanasiou, I., Rehman, I., & Lee, K. Y. (2023). School-based cognitive-behavioural therapy for children and adolescents with social anxiety disorder and social anxiety symptoms: A systematic review. *PLoS ONE, 18*. https://doi.org/10.1371/journal.pone.0283329
Weber, J., Heming, M., Apolinário-Hagen, J., Liszio, S., & Angerer, P. (2024). Comparison of the perceived stress reactivity scale with physiological and self-reported stress responses during ecological momentary assessment and during participation in a virtual reality version of the trier social stress test. *Biological Psychology*. , Article 108762. https://doi.org/10.1016/j.biopsycho.2024.108762
Wechsler, T. F., Kümpers, F., & Mühlberger, A. (2019). Inferiority or even superiority of virtual reality exposure therapy in phobias?-A systematic review and quantitative meta-analysis on randomized controlled trials specifically comparing the efficacy of virtual reality exposure to gold standard *in vivo* exposure in agoraphobia, specific phobia, and social phobia. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.01758
Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments, 7*(3), 225–240. https://doi.org/10.1162/105474698565686
Wong Sarver, N., Beidel, D., & Spitalnick, J. (2014). The feasibility and acceptability of virtual environments in the treatment of childhood social anxiety disorder. *Journal of Clinical Child and Adolescent Psychology, 43*(1), 63–73. https://doi.org/10.1080/15374416.2013.843461

Mathieu Delangle, **PhD,** is an engineer with a doctorate in mechanical engineering from Centrale Nantes, France, specializing in virtual reality and ergonomics. His research focuses on digital technologies supporting human factors, design of inclusive environments, and the influence of VR environments on cognition, behavior, and physiological responses in children and adults with functional and/or language-specific needs.

Anne Moïse-Richard, M.P.O., is a Speech-Language Pathologist at the Marie-Enfant Rehabilitation Center (CHU Sainte-Justine, Montreal) and lecturer at the Université du Québec à Trois-Rivières. Her clinical and research interests concern childhood stuttering, social communication, and virtual reality–based interventions.

Anne-Lise Leclercq, PhD, is an Associate Professor of Speech-Language Pathology at the University of Liège, Belgium. Her work focuses on developmental speech and language disorders, fluency disorders, and the use of digital tools in clinical practice.

David Labbé, PhD, is a Full Professor in the Department of Software and IT Engineering at the École de technologie supérieure (ÉTS) in Montreal, Canada. His research focuses on virtual and augmented reality applications in human motion analysis, motor control, and orthopedic rehabilitation. He is also a researcher at the Centre de recherche du CHUM (CRCHUM), affiliated with the University of Montreal.

Stéphane Bouchard, PhD, is a Full Professor in the Department of Psychoeducation and Psychology at the Université du Québec en Outaouais (UQO), where he holds the Canada Research Chair in Clinical Cyberpsychology. He is also the co-director of the UQO Cyberpsychology Laboratory. His research focuses on the clinical effectiveness of virtual reality environments for treating mental health disorders, particularly anxiety and phobias.

Sheldon Andrews, PhD, is an Associate Professor in the Department of Software and IT Engineering at the École de technologie supérieure (ÉTS) in Montreal, Canada. His research interests include real-time physics simulation, 3D character animation, motion capture and measurement-based modeling for virtual environments.

Lucie Ménard, PhD, is a Full Professor of Linguistics at the Université du Québec à Montréal and Director of the Phonetics Laboratory. Her research examines speech development, audiovisual integration, and prosody in Quebec French, with applications in speech synthesis and virtual reality.